SPREADING PROCESSES IN HUMAN SYSTEMS

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

im Fach *Physik*
in der Spezialisierung *Theoretische Physik*

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät der
Humboldt-Universität zu Berlin

von

BENJAMIN F. MAIER, M. Sc.

Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät
Prof. Dr. Elmar Kulke

Gutachter
1. Prof. Dr. Igor Sokolov
2. Prof. Dr. Dirk Brockmann
3. Prof. Dr. Sune Lehmann

Tag der mündlichen Prüfung: 16.12.2019

The cover art displays a modified random geometric graph (see Sec. 2.2.3) with nodes sampled according to the density of an edited depiction of a human head (the original artwork was created by Juan Diego Lugo and was edited with permission; for the network sampling procedure see the corresponding code sample by Vedran Sekara and Ulf Aslak at github.com/ulfaslak/network_yo_face). A teal colored path follows the steps of a single random walker (see Sec. 3.1). The pink tree corresponds to the infection tree of an SI process seeded at a single node (an SIS process with vanishing recovery rate, see Sec. 3.2).

Abstract

Human systems have been modeled and analyzed on the basis of complex networks theory in recent time. This abstraction allows for thorough quantitative analyses to investigate which structural and temporal features of a system influence the evolution of spreading processes, such as the passage of information or of infectious diseases. Key research questions are related to the prediction of signal passage times and the systemic susceptibility for disease. The first part of this work investigates how the ubiquitous modular hierarchical structure of static real-world networks allows for fast delivery of messages. Based on prior research concerning the informed passage of messages in networks with small-world features, it is shown here that a variety of modular hierarchical network models are of optimal structure for fast delivery of signals, even if the spreading process follows purely random dynamics. To this end, new heuristics are developed to evaluate random walk mean first passage times and cover times on locally clustered networks. A comparison to average medium approximations demonstrates that the emergence of these minima are pure network phenomena. It is further found that not all modular hierarchical network models provide optimal message delivery structure. In the second part, temporally varying face-to-face contact networks are investigated for their susceptibility to infection. Several studies have shown that people tend to spend time in small, densely-connected groups or in isolation, and that their connection behavior follows a circadian rhythm. To what extent both of these features influence the spread of diseases is as yet unclear. To this end, a new temporal network model is devised here which allows for a thorough investigation of a variety of network properties and how they affect disease spreading. Based on this model, circadially varying networks are interpreted as moving on trajectories through a newly defined state space with one structural and one temporal dimension. It is further revealed that in many temporally varying networks the system becomes less susceptible to infection when the time-scale of the disease approaches the time-scale of the network variation. This result is in direct conflict with findings of other studies which predict increasing susceptibility of temporal networks, a discrepancy which is attributed to the invalidity of a widely applied approximation. Therefore, the results presented here imply that new theoretical advances are necessary to study the spread of diseases in temporally varying networks.

Keywords

complex networks, random walks, epidemic spreading, modular hierarchical networks, small-world effect, cover time, first passage time, temporal networks, epidemic threshold, face-to-face networks, circadian human activity

Zusammenfassung

Menschliche Systeme werden seit einiger Zeit auf der Basis der Komplexe-Netzwerk-Theorie modelliert und analysiert. Diese Abstraktion ermöglicht es quantitativ zu analysieren, welche strukturellen und zeitlichen Eigenschaften eines Systems Ausbreitungsprozesse beeinflussen, z.B. von Informationen oder von infektiösen Erregern. Damit verbundene Kernfragen der Forschung beschäftigen sich mit der Vorhersage von Ankunftszeiten in der Nachrichtenverbreitung und der Anfälligkeit eines Systems für Krankheiten. Der erste Teil dieser Arbeit untersucht, wie die omnipräsente modular-hierarchische Struktur statischer, real existierender Netzwerke die schnelle Verbreitung von Signalen ermöglicht. Basierend auf früherer Forschung über informierte Nachrichtenverbreitung in "Small-World"-Netzwerken wird hier gezeigt, dass eine Vielzahl von modular-hierarchischen Netzwerkmodellen eine für schnelle Verbreitung optimale Struktur besitzen, auch wenn diese lediglich einer rein zufälligen Dynamik folgt. Zu diesem Zweck werden neue Heuristiken entwickelt um die Random-Walk-Observablen "First Passage Time" und "Cover Time" auf lokal geclusterten Netzwerken zu ermitteln. Vergleiche mit der Approximation eines gemittelten Mediums zeigen, dass das Auftreten der beobachteten Minima dieser Observablen ein reiner Netzwerkeffekt ist. Es wird weiterhin dargelegt, dass nicht alle modular-hierarchischen Netzwerkmodelle dieses Phänomen aufweisen. Im zweiten Teil der Arbeit werden zeitlich veränderliche Face-to-Face Kontaktnetzwerke in Hinblick auf ihre Anfälligkeit für Krankheitsausbreitung untersucht. Mehrere Studien belegen, dass Menschen vornehmlich Zeit in Isolation oder kleinen, stark verbundenen Gruppen verbringen, und dass ihre Kontaktaktivität einem zirkadianen Rhythmus folgt. Auf welche Weise diese Eigenschaften die Ausbreitung einer Krankheit beeinflussen ist noch unergründet. Daher wird hier ein neues temporales Netzwerkmodell entwickelt, anhand dessen spezifische Netzwerkeigenschaften und ihr Einfluss auf Krankheitsausbreitung untersucht werden. Basierend auf diesem Modell wird gezeigt, dass zirkadian variierende Netzwerke Trajektorien folgen in einem neu definierten Zustandsraum mit einer strukturellen und einer zeitlichen Dimension. Es wird weiterhin dargelegt, dass viele zeitlich veränderliche Netzwerke weniger anfällig für Epidemien sind, wenn sich die Zeitskala der Epidemie der Zeitskala der Netzwerkveränderungen annähert. Dieses Ergebnis steht in direktem Konflikt mit den Ergebnissen anderer Studien, die eine erhöhte Anfälligkeit der Systeme vorhersagen, eine Diskrepanz, die auf die Invalidität einer oft verwendeten Approximation zurückgeführt wird. Aus diesem Grund legen die Ergebnisse dieser Arbeit nahe, dass die Entwicklung neuer theoretischer Methoden notwendig ist, um Krankheitsausbreitungen in zeitlich veränderlichen Netzwerken zu untersuchen.

Schlagwörter

Komplexe Netzwerke, Random Walks, epidemische Ausbreitung, modular-hierarchische Netzwerke, Small-World-Effekt, Cover Time, First Passage Time, zeitlich veränderliche Netzwerke, Face-to-Face-Netzwerke, zirkadiane menschliche Aktivität

Contents

1 Introduction **21**

I Human Systems as Networks: Theory & Data **29**

2 Networks **31**

2.1 Network Properties 31
2.1.1 The Adjacency Matrix and Some Other Network Definitions 31
2.1.2 Node Degree and Degree Distribution 33
2.1.3 A Word on Heavy-Tailed Distributions 35
2.1.4 Clustering Coefficient 36
2.1.5 Networks Composed of Multiple Components . 38
2.1.6 Shortest Path Lengths 39
2.1.7 Modular Hierarchical Structure 40
2.2 Static Network Models 43
2.2.1 Regular Networks and Lattices 43
2.2.2 Erdős–Rényi Random Networks 45
2.2.3 Random Geometric Graphs 47
2.2.4 Small-World Effect and Corresponding Network Models 47
2.2.5 Barabási-Albert Networks 51
2.3 Temporal Networks 52
2.3.1 Temporal Contact Networks 52
2.3.2 Temporal Network Observables 56
2.3.3 Empirical Temporal Human Proximity Data Sets 58
2.4 Temporal Network Models 64
2.4.1 Memory-Kernel Models 64
2.4.2 Edge Activity Model 67
2.4.3 Other Models 68

3 Spreading Dynamics **69**

3.1 Discrete-Time Random Walks 69
3.1.1 Random Walks on Networks 70
3.1.2 Graph Laplacian 71
3.1.3 Mixing Time 72
3.1.4 (Mean) First Passage Time 73
3.1.5 Global Mean- and Pair-Averaged First Passage Time 74

3.1.6 Cover Time 76
3.2 The SIS Contagion Model 76
3.2.1 Mean-Field SIS and the Epidemic Threshold . . 77
3.2.2 Static Networks and the Individual-Based Markov-Chain Approximation 79
3.2.3 SIS Simulation Algorithms on Static Networks . 82
3.2.4 SIS on Temporal Networks 86

II Random Walks on Modular Hierarchical and Other Small-World Networks 91

4 Small Worlds Are All the Same 93
4.1 Self-Similar Modular Hierarchical Random Networks . 94
4.1.1 Model Definition and Structural Phases 94
4.1.2 Network Properties 96
4.2 Modified Watts–Strogatz Small-World Networks 98
4.2.1 Structural Definition and Sampling Algorithm . 98
4.2.2 Network properties 100
4.3 Power-Law Small-World (PLSW) Networks 102
4.3.1 General Construction of PLSW Networks with Constant Mean Degree 103
4.3.2 Application to 2-Torus and Square 105
4.3.3 Application to One-Dimensional Lattice PLSW Networks 107
4.4 Structural Relation of Small-World Models 108
4.4.1 Mapping SSMH Networks to One-Dimensional Lattice PLSW Networks 108
4.4.2 Mapping modified Watts–Strogatz Networks to One-Dimensional Lattice PLSW Networks . . . 109
4.4.3 All Small-World Network Models Appear to be of Modular Hierarchical Structure 112
4.5 Summary and Discussion 114

5 Passage Time Statistics for Locally Clustered Networks 117
5.1 Average Medium Approximation 118
5.1.1 SSMH First Passage Time Distribution and Pair-Averaged First Passage Time 119
5.1.2 Mixing Time and Small-World Effect in the Modified Watts–Strogatz Model 122
5.2 Actual Network Realizations 124
5.2.1 Numerical Results for Pair-Averaged First Passage Time 124
5.2.2 Heuristic Explanation of Minimum 125
5.3 Summary and Discussion 129

6 Cover Time for Arbitrary Complex Networks 131
6.1 Heuristic Evaluation Method 131
6.1.1 Derivation 131
6.1.2 Cover Time of Networks With Equal GMFPTs . 134

6.2 Numerical Results . . . 135
6.3 Error Analysis . . . 136
6.3.1 Low-Dimensionality Induced Deviations . . . 136
6.3.2 Systematic Error: Continuous-Time Approximation . . . 137
6.3.3 Systematic Error: Approximation of Mean Cover Time Integral . . . 139
6.4 Cover Time on Modular Hierarchical Small-World Models . . . 141
6.4.1 Average Medium Approximation . . . 141
6.4.2 Actual Network Realizations . . . 142
6.5 Summary and Discussion . . . 144

III Epidemics on Temporal Face-to-Face Contact Networks 145

7 Flockworks: Analyzing Temporal Face-to-Face Contact Networks With a Simple Model 147
7.1 The Poissonian Flockwork Model . . . 148
7.1.1 Model Definition . . . 148
7.1.2 Group Sizes and Equilibrium State . . . 149
7.1.3 Node Degrees . . . 153
7.1.4 Distributions of (Inter-)Contact Durations and Group Life-Times . . . 155
7.1.5 Comparing the Basic Poissonian Model to Real Data . . . 156
7.2 Real-World Systems as Based on the Flockwork Model 158
7.2.1 Inferring Time-Varying Activity Rates and Network State Variables . . . 158
7.2.2 Application to Real-World Face-to-Face Systems 162
7.3 Summary and Discussion . . . 171

8 The SIS-Model on Temporal Contact Networks 175
8.1 Introduction . . . 175
8.2 Spreading on Temporal Network Models . . . 177
8.2.1 General Considerations . . . 177
8.2.2 Epidemic Response . . . 179
8.2.3 Epidemic Threshold . . . 184
8.2.4 The Case of Infinite Infectiousness . . . 187
8.3 Spreading on Real-World Temporal Networks . . . 190
8.3.1 Methods . . . 190
8.3.2 Averaged Networks . . . 191
8.3.3 The Influence of Circadian Activity . . . 191
8.3.4 Comparison of Gillespie Simulations and Markov-Chain Approximation . . . 194
8.3.5 The Influence of Rebinning Temporal Network Data . . . 196
8.3.6 Surrogate Networks . . . 198
8.4 Summary and Discussion . . . 199

9 Summary and Conclusions **205**

10 Acknowledgments **211**

Bibliography **215**

A Infomap: A Flow-based Algorithm to Find Modular Hierarchical Network Partitions **227**

B Evaluations Concerning Small-World Models **231**
B.1 SSMH Clustering Coefficient 231
B.2 PLSW Networks 232
B.2.1 Extension to a Finite Set of Distances 233
B.2.2 Extension to Categorical Distances 234
B.3 Average Medium Approximations 235
B.3.1 Time Evolution of Walker Distribution on SSMH Model 235
B.3.2 Global Mean First Passage Time 236

C Additional Flockwork Evaluations **239**
C.1 Group Evolution 239
C.2 Efficient Flockwork Equilibrium Configuration Sampling Algorithm 246
C.3 Link Decay Rate 250
C.4 Testing Rate-Inference with Synthetic Time-Varying Flockwork Data 250
C.5 Group Size Evolution of Infected for Infinite Infectiousness 253

D Gillespie's Stochastic Simulation Algorithm **257**
D.1 Homogeneous Poisson Processes 257
D.1.1 A Single Event 257
D.1.2 Multiple Events 258
D.2 Inhomogeneous Poisson Processes 260
D.2.1 A Single Event 260
D.2.2 Multiple Events 262
D.2.3 Application to Locally Constant Rates 264

E SIS-Simulations on Temporal Networks **267**
E.1 Finite Infection Rates 267
E.2 Infinite Infection Rate 268
E.3 Fitting Epidemic Response Curves 269
E.4 Simulations on Real-World Data 270
E.4.1 Methods 270
E.4.2 Circadian Influence Analysis 272
E.5 SIS-Model on a Pair 272
E.5.1 The ε-SIS-Model and the MCA 272
E.5.2 Exact Evaluation 277
E.5.3 Epidemic Threshold 278

F Rate and State-Space Trajectory Inference from Real Data **283**

List of Publications

The following articles were published based on the results presented in Part II of this dissertation:

1. Maier, B. F. and Brockmann, D. "Cover time for random walks on arbitrary complex networks". *Physical Review E* 96.4 (Oct. 2017), p. 042307.

2. Maier, B. F., Huepe, C., and Brockmann, D. "Modular hierarchical and power-law small-world networks bear structural optima for minimal first passage times and cover time". *Journal of Complex Networks*, cnz010 (Apr. 2019).

3. Maier, B. F. "Generalization of the small-world effect on a model approaching the Erdős–Rényi random graph". *Scientific Reports* 9.1 (June 2019), p. 9268.

Most of the network illustrations in this work were made using a newly developed network visualization package which is currently under review:

4. Aslak, U. and Maier, B. F. "Netwulf: Interactive visualization of networks in Python". *Journal of Open Source Software* (under review).

List of Figures

1.1 The author's Facebook friend network. 23

2.1 Example networks and their respective adjacency matrices. 33
2.2 A small network and some node neighborhoods. 34
2.3 Degree distributions of exemplary networks. 35
2.4 Different model distributions to illustrate heavy tails. 36
2.5 Illustration of the clustering coefficient. 37
2.6 Component size distribution of the Facebook friend network. 38
2.7 An illustration of a shortest path. 39
2.8 Illustration of how networks are analyzed for their modular hierarchical structure in this thesis. 40
2.9 Modular hierarchical structure of various social networks. 41
2.10 Modular hierarchical structure of the Western U.S. power grid. 43
2.11 A two-dimensional square lattice of finite side length. 43
2.12 A k-regular nearest-neighbor lattice 43
2.13 Evaluating the triangles per node in the k-regular nearest neighbor lattice. 44
2.14 The complete network. 44
2.15 A sample Erdős–Rényi network and its degree distribution. 46
2.16 Largest component and component size distribution on samples of the ER network. 46
2.17 A sample random geometric graph. 47
2.18 Watts–Strogatz networks. 48
2.19 Properties of Watts–Strogatz networks 49
2.20 Construction illustration for Kleinberg networks. 49
2.21 Watts–Dodds–Newman modular hierarchical networks. 50
2.22 A sample Barabási-Albert network and its degree distribution. 51
2.23 Illustration of an example temporal network. 54
2.24 Discretizing time by binning and sampling a temporal network. 55
2.25 The DTU temporal network data set. 60
2.26 The HT09 temporal network data set. 62
2.27 The HS13 temporal network data set. 63
2.28 Comparison of structural and temporal statistics across the DTU, HT09, and HS13 data sets and the influence of binning. 65

3.1 Illustration of a single time step in a random walk process on an undirected unweighted network. 70
3.2 Illustration of the random walk transition matrix as a weighted directed network and the transition matrix with a sink node. 71

3.3 Schematic representation of response curve and epidemic threshold in the SIS system. 79
3.4 Results of the Markov-chain approximation regarding epidemic response curves on static network models. 81
3.5 Gillespie's algorithm for SIS dynamics on static networks. 82
3.6 Epidemic response curves on the largest component of the author's Facebook friend network, comparing multiple evaluation methods. 84
3.7 Epidemic response curves on model networks, obtained via Gillespie simulations and the individual-based Markov-chain approximation. 85
3.8 Epidemic response curves on the rebinned HT09 data set according to Valdano *et al.* 86
3.9 Rescaled epidemic response curves on the HT09 data set according to Valdano *et al.* 86
3.10 Epidemic response curves on the rebinned HT09 data set according to Speidel *et al.* 88

4.1 Schematic hierarchical modular structure of the SSMH model. 94
4.2 The SSMH modular hierarchical phases. 96
4.3 Properties of the SSMH network model for increasing control parameter. 97
4.4 Schematic representation of the modified small-world model. 99
4.5 Network samples from the modified Watts–Strogatz model for increasing structural control parameter. 100
4.6 Schematic figure for the evaluation of the areas of summation to find the expected number of triangles per node in the modified Watts–Strogatz model. 101
4.7 Network properties of the modified Watts-Strogatz model. 101
4.8 How to construct the connection probability ro achieve constant mean degree in power-law small-world (PLSW) models. 104
4.9 Sampling procedure for the unit square PLSW model. 106
4.10 Network properties of PLSW small-world networks with underlying topology of the unit square and the 2-torus. 107
4.11 Properties of the one-dimensional lattice PLSW network model. 108
4.12 Method for redistributing the probability mass function of the one-dimensional lattice PLSW network model. 109
4.13 Comparison of the connection probability functions associated with SSMH networks and with one-dimensional lattice PLSW networks. 110
4.14 Topological phases of three small-world models as a function of their respective structural control parameters. 111
4.15 Node properties of the SSMH, the modified WS and the one-dimensional lattice PLSW network model. 112
4.16 Results of the modular hierarchical partition detection on small-world network models using *Infomap*. 113

5.1 Time evolution of the random walker density at hierarchical distance from the sink for the average medium approximation of the SSMH model. 118
5.2 The SSMH average medium arrival time cdf. 120
5.3 Pair-averaged FPT for the SSMH model and the 1D lattice model computed from simulations. 121
5.4 The SSMH average medium arrival time pmf 121
5.5 The small-world effect as illustrated by the modified Watts-Strogatz clustering coefficient and mixing time. 123
5.6 Pair-averaged first passage time computed numerically for the SSMH and the 1D lattice PLSW models. 124
5.7 Pair-averaged FPT for network samples of the main small-world models introduced in the previuos chapter. 125
5.8 Illustration of the heuristic method to find the global mean first passage time (FPT) of a sink node with non-zero clustering. 126
5.9 Heuristic result for the pair-averaged FPT. 127
5.10 Comparison between the actural pair-averaged FPT and the heuristic lower bound on a 2-torus PLSW model. 128

6.1 Illustration of the heuristic approach to compute the mean cover time. 132
6.2 The mean cover time of the largest component of Erdős-Rényi, Barabási-Albert, and k-regular networks. 133
6.3 The heuristic cover time evaluation method yielding results with rather large deviations from simulations on lattices. 137
6.4 Cover time pdfs and ccdfs for four example networks. 138
6.5 Mean cover times on the largest component of various real-world networks. 139
6.6 The relative error of the mean cover time heuristic is increasing with increasing mixing time per node. 140
6.7 Relative error between two integrals computing the mean cover time. 140
6.8 The mean cover time as estimated with an average medium approximation of SSMH and 1D lattice PLSW networks. 142
6.9 The mean cover time on actual network realizations of the (modular hierarchical) small-world models. 143
6.10 Relative errors between cover time simulation results and heuristic results for the small-world models. 144

7.1 The possible events of a Flockwork model. 149
7.2 Flockwork equilibrium configurations for increasing reconnection probability 150
7.3 The temporal evolution of the Flockwork model. 151
7.4 Configuration examples following the temporal evolution of the Flockwork model. 152
7.5 Flockwork equilibrium group-size distribution for increasing reconnection probability and number of nodes. 153

7.6 Mean group size, mean number of components and second moment of the equilibrium group-size distribution compared to asymptotics. 154
7.7 Flockwork equilibrium mean degree as a function of the reconnection probability and number of nodes, compared to the asymptotics. 154
7.8 Flockwork equilibrium degree distributions for increasing reconnection probability and number of nodes. 155
7.9 Ergodic group-size and life-time distributions of single Flockwork simulations. 157
7.10 Example network behavior for different positions in the Flockwork state space. 161
7.11 Inferred time-varying rates and expected mean degree of the real-world data sets DTU, HT09 and HS13. 163
7.12 Inferred state-space trajectories of the real-world data sets DTU, HT09 and HS13. 164
7.13 Structural and statistical comparison between the DTU data set and a corresponding time-varying Flockwork surrogate. 166
7.14 Structural and statistical comparison between the HT09 data set and a corresponding time-varying Flockwork surrogate. 167
7.15 Structural and statistical comparison between the HS13 data set and a corresponding time-varying Flockwork surrogate. 169
7.16 The effect of binning on measuring groups. 170
7.17 Comparison of binned Flockwork group life-time distributions. 171

8.1 Schematic representation of response curve and epidemic threshold. 175
8.2 Endemic state on the Flockwork model obtained by Gillespie SIS simulations as a function of average mean degree and contact-renewal number. 180
8.3 Endemic state on the edge activity model obtained by Gillespie SIS simulations as a function of average mean degree and contact-renewal number. 181
8.4 Ratio of infected in equilibrium for Gillespie SIS simulations and the Markov-chain approximation on the Flockwork model. 182
8.5 Ratio of infected in equilibrium for Gillespie SIS simulations and the Markov-chain approximation on the edge activity model. 183
8.6 Non-zero ratio of infected in equilibrium as rescaled with the fit parameters. 185
8.7 Epidemic threshold extracted from Gillespie SIS simulations using fits. 186
8.8 Epidemic threshold extracted from the individual-based Markov-chain approximation using fits. 187
8.9 Endemic state on the Flockwork model for an infinitely infectious disease. 189
8.10 Epidemic response curves from Gillespie simulations on the averaged temporal networks and the temporally resolved networks for a low recovery rate. 191

8.11 The influence of circadian activity on epidemic response curves and epidemic thresholds. 192
8.12 Circadian temporal network trajectory crossing the critical surface to reach the absorbing state. 193
8.13 Mean endemic state and minimal endemic state for varying recovery rates on the DTU data. Comparison between Gillespie simulations and results from the Markov-chain approximation. 194
8.14 Epidemic response curves for DTU, HT09, and HS13, comparing results from Gillespie simulations to those of the Markov-chain approximation. 195
8.15 Epidemic response curves for DTU, HT09, and HS13, comparing results from Gillespie simulations on both original and rebinned data. 196
8.16 Epidemic response curves for rebinned data of DTU, HT09, and HS13, comparing results from Gillespie simulations to those of the Markov-chain approximation. 197
8.17 Epidemic response curves for Gillespie simulations on Flockwork surrogate data generated to resemble the single-snapshot structure and circadian activity variation of the DTU data set, compared to simulation results on the original data. 198
8.18 Epidemic response curves for surrogate data generated from the Flockwork and edge activity models to mimick the DTU data set and compared to the result of the corresponding Markov-chain approximation. 198
8.19 Epidemic response curves for surrogate data generated from the Flockwork model for the DTU data set, both from simulations and the individual-based Markov-chain approximation. 199

B.1 Pairwise distance distributions of points uniformly distributed in a square and a 2-torus. 232
B.2 Connection probability and distance distribution of node pairs in the unit square and 2-torus PLSW models. 233

C.1 The distribution of numbers n_g of g-sized groups in the equilibrium state. 248
C.2 Group size distribution as a mean over independently constructed equilibrium configurations using all variants of the proposed sampling algorithm. 249
C.3 Measurements vs. Eq. ((7.13)). Each data point is a mean over 10 independent simulations of runtime $T = 2 \times 10^6 (N\gamma)^{-1}$. 250
C.4 Inferring time-varying rates and the state-space trajectory of a corresponding Flockwork simulation. 252

E.1 Temporal evolution of ratio of infected in an infinitely infectious SIS process. 269
E.2 Normalized residuals of the fits to the SIS simulation data on the Flockwork model. 270
E.3 Normalized residuals of the fits to the SIS simulation data on the edge activity model. 271

E.4 The influence of circadian activity on epidemic response curves and epidemic thresholds for both exact simulations and the Markov-chain approximation on real-world temporal networks. 273
E.5 Enlarged illustration of the influence of circadian activity on epidemic thresholds, as a comparison between Gillespie simulations and the individual-based Markov-chain approximation on real-world networks. 274
E.6 The influence of circadian activity on epidemic response curves and epidemic thresholds for both exact simulations and the Markov-chain approximation on Flockwork surrogate networks. 275
E.7 Enlarged illustration of the influence of circadian activity on epidemic thresholds, as a comparison between Gillespie simulations and the individual-based Markov-chain approximation on Flockwork surrogate networks. 276
E.8 The ε-SIS model on a single pair of nodes. 277

F.1 Inference of the state-space trajectory from an artifical varying-rate Flockwork simulation. 283
F.2 Inference of the reconnection rate, disconnection rate, and average mean degree from an artifical varying-rate Flockwork simulation. 284
F.3 Inference of the reconnection rate from the DTU data. 285
F.4 Inference of the expected mean degree from the DTU data. 286
F.5 Inferred trajectory in the state space from the DTU data. 287
F.6 Inference of the reconnection rate from the HT09 data. 288
F.7 Inference of the expected mean degree from the HT09 data. 289
F.8 Inferred trajectory in the state space from the HT09 data. 290
F.9 Inference of the reconnection rate from the HS13 data. 291
F.10 Inference of the expected mean degree from the HS13 data. 292
F.11 Inferred trajectory in the state space from the HS13 data. 293

List of Tables

2.1 Network properties of a collection of real-world networks and relatedness to modular hierarchical clustering. 42

6.1 Mean cover times of simple discrete-time random walks on the largest component of various real-world networks. 135
6.2 Mean cover times of simple discrete-time random walks on subway networks of big cities. 136

E.1 Chosen values of the equilibration time t_{eq} for SIS simulations on both the Flockwork model as well as the edge activity model. 268

Nomenclature

α	Active reconnection rate per node
β	Active disconnection rate per node
$\boldsymbol{A}$	Adjacency matrix
$\boldsymbol{L}$	Unnormalized graph Laplacian
$\boldsymbol{W}$	Transition matrix
η	Infection rate per SI-link
$\Gamma(x)$	The Gamma function
γ	Disconnection rate per node
λ_2	The unnormalized graph Laplacian's second smallest eigenvalue
$\langle c \rangle$	Mean number of components
$\langle e \rangle$	Mean number of edges in the network
$\langle g \rangle$	Mean group size
$\langle k \rangle$	Mean degree
$\mathcal{B}$	Binomial distribution
$\mathcal{E}$	Exponential distribution
$\mathrm{CV}[X]$	Coefficient of variation of random variable X
$\mathrm{Nei}(u)$	Neighborhood of node u
$\mathrm{Var}[X]$	Variance of random variable X
ω^+	Link activation rate
ω^-	Link deactivation rate
$\overline{\langle k \rangle}$, k_0	Temporally averaged mean degree
ρ	Recovery rate per infected node
τ_C	Contact duration (duration of a single link)
τ_{IC}	Inter-contact duration (duration of a single link)
C	Global clustering coefficient

$I^\star$	Number of infected in equilibrium
$i^\star$	Ratio of infected in equilibrium
m, e	The total number of edges in a network
N	Number of nodes in a network
n_g	Number of groups of size g in a network
P, p	Probabilities (depending on context)
R_0	Mean-field basic reproduction number
T	Part II: cover time, Part III: observation period of a temporal network
t_{mix}	Random walk mixing time (inverse of the transition matrix's second largest eigenvalue ω_{N-1})
AMA	Average medium approximation
ccdf	complementary cumulative distribution function
cdf	cumulative distribution function
ER	Erdős–Rényi
FPT	First passage time
GMFPT	Global mean first passage time
MCA	Markov-chain approximation
MFPT	Mean first passage time
pdf	probability distribution function
PLSW	Power-law small-world
pmf	probability mass function
RGG	Random geometric graph
SSMH	Self-similar modular hierarchical
WS	Watts–Strogatz

1
Introduction

If you have spotted this dissertation in the 'Statistical Physics' section of your library, raised an eyebrow, and opened it wondering why this is placed here, then please be assured that such a reaction is not outlandish. This work is both titled "Spreading Processes in Human Systems" and located in the Physics section because it is rooted in complex network theory, an inter-disciplinary field of strong overlap with Statistical Physics and the study of Dynamic Systems. For the last century, researchers, including physicists and social scientists, have been developing methods to model and analyze human systems as *networks*. In these, collection of entities (*nodes*) are connected to each other via *links*. In such abstractions, nodes may represent a variety of things including for instance transformer stations in human-built power grids [1], pages in the world wide web [2], airports [3, 4], or people [1, 5], with links corresponding to high-voltage transmission lines, hyper-links, airline connections, or social contacts. An example *social network*, which refers to a collection of individuals that are linked via social relationships, is displayed in Fig. 1.1—it is the author's 'Facebook' friend network[1] of the year 2014 [6]. This network will serve an illustratory purpose in this introduction and in Part I of this dissertation. In Fig. 1.1, nodes are drawn as disks and represent individuals the author acquainted. Links between any pair of nodes, illustrated as grey lines, mean that those two nodes actively decided to be 'Facebook friends' with each other, i.e. be connected on the platform such that they can exchange information, e.g. via messages. Often, the purpose of describing a system as a network is connected to questions about how information spreads in them: How long does it take for a message to travel from one entity to another? How severe will a spreading disease be for a system? These rather abstract questions can be connected to applications and effects like efficient routing in wireless local area networks [7], the small-world effect [1, 8], synchronization in oscillator systems [9, pp. 138 ff.], computing the two-point resistance of a resistor network [10], predicting the arrival time of a pandemic disease [3, 4], and the development of monitoring and vaccination strategies to mitigate the spread of an infection [11] and [9, pp. 207 ff.]. In order for such complex applications to be fruitful, it is important to first understand the basic principles of the emergence of complex net-

[1] In the case of a reader luckily not being familiar with Facebook, or in the other lucky case that Facebook has vanished by the time this is read: Facebook is/was an online social media platform where users can mutually connect to exchange information.

works, their structural properties, and the influence these properties have on primitive dynamics.

In this thesis, theoretical methods to investigate the influence of basic structural properties of complex networks on the outcome of simple dynamic systems are analyzed and further developed. In the following, current central problems of the field are introduced, based on which the argumentation and structure of the remaining work will be laid out.

A variety of empirically collected networks from different contexts display a similar structural behavior: Nodes can be arranged in densely[2] connected modules, which may consist of various submodules and so forth, recursively building a hierarchical structure of communities. Taking the author's Facebook network in Fig. 1.1 as an example, nodes can be grouped according to certain contexts in which the author acquainted the corresponding individuals, e.g. at Humboldt University of Berlin (HU Berlin). Nodes in this community have more connections to other nodes of the same community than they have to nodes of other communities. Within the largest community of HU Berlin, however, there are smaller communities which might correspond to certain classes the individuals took together, thus building a community within a community. While this kind of structure is often associated with social networks [12–14] it is similarly ubiquitous in other systems, such as networks of neurons in the brain [15–21], cellular and metabolic networks [14, 22, 23], human transport networks [24, 25], and ecological systems [14, 26, 27]. Many studies concentrated on developing methods to infer such hierarchical structures [14, 22, 24, 28, 29]. Less research focused on the question why such structures emerge and persist in the first place [13, 25, 26, 30]. One hypothesis is that such structure is optimal for certain dynamic processes such as for fast passage of information or for the stability of an ecological system [26]. Based on this assumption, two particular studies revealed that a class of network models called 'small-world' networks are optimal for efficient message spreading [13, 31]. In such small-world modeling approaches, it is usually assumed that nodes are part of a strongly connected, clustered environment but possess a small number of 'long-range' connections to other parts of the network, where a single *control parameter* regulates the amount of those long-range connections, interpolating between two structural limits: a localized clustered limit and a randomized limit [1, 13, 31]. Such networks can be of optimal structure to enable fast message forwarding under the condition that acting nodes have access to local information about the network and the spreading process itself, in which case they are able to identify short paths between a source and a target [13, 31]. In this context, a modular hierarchical network model was devised, showing that in between the described structural limits, a modular hierarchical structure is optimal for such informed message spreading (or informed searching, which is equivalent) [13]. The analysis was based on an approximate network description, only analyzing average links instead of actual

[2] In a dense network, many of all node pairs are connected by links. Network density is discussed in Sec. 2.1.2.

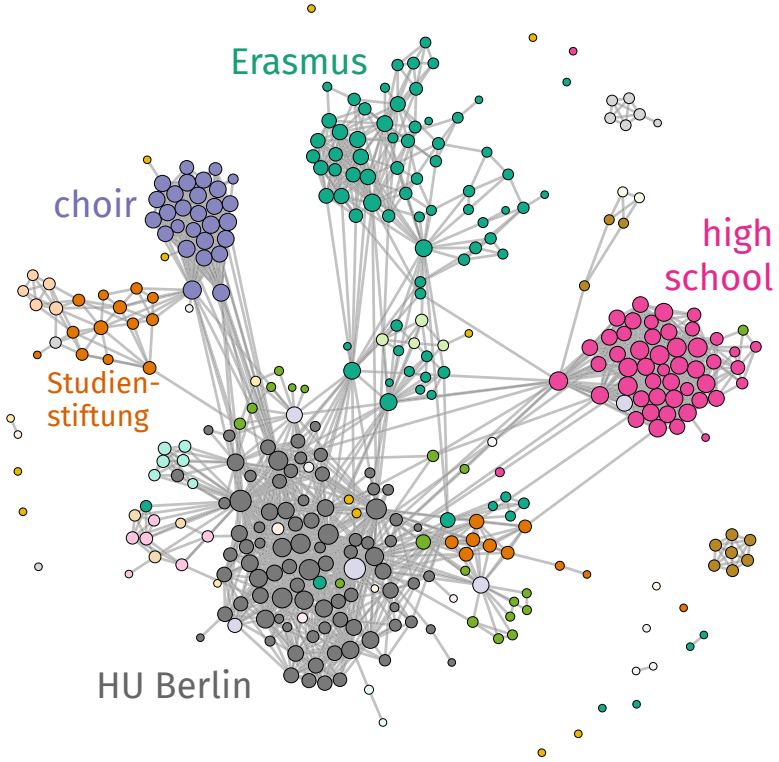


Figure 1.1: The author's Facebook friend network. Each node (depicted by a colored disk) represents a person the author was "friends" with on the online social media platform 'Facebook' in 2014 [6]. A link (straight line) between two nodes implies that both nodes are "Facebook friends", which means they both actively accepted the social media connection—they were therefore able to see each other's 'feeds' and write messages to each other, i.e. able to exchange information. The author, who would be connected to each of the nodes displayed in this figure, was deleted from the network. Nodes are colored according to the main real-world context through which the author acquainted them, manually assigned. The largest contexts "HU Berlin", "high school", "Erasmus", "choir", and "Studienstiftung" all form strong clusters in which many nodes have contacts to each other. The node size varies according to the number of connections it has to other nodes in the network: A larger node has more connections than a smaller node.

network structures [32, pp. 723–729]. A variety of natural processes, however, might not follow the logics of informed searches but have more random, diffusion-like properties [33, 34]. Also, links of networks rooted in the real world are not necessarily correctly reflected by average values but might follow certain probability distributions. Therefore, in Part II of this thesis I investigate the hypothesis that modular hierarchical networks emerge because they bear structural optima for efficient random i.e. maximally uninformed spreading of signals, modeled as random walks here. The argumentation will be laid out as follows.

In Sec. 2.1, networks are formally defined and relevant network properties are summarized. This includes the clustering coefficient, the degree distribution, and a definition of modular hierarchical structure. Subsequently, small-world network models including a modular hierarchical network model are summarized in Sec. 2.2. While those network models share certain properties, it is yet unclear how to properly map their control parameters in a way that their structures correspond. The random walk spreading process is introduced in Sec. 3.1, including its temporal observables 'pair-averaged first passage time' and 'cover time': the former measures how long it takes for a random signal to travel between any two nodes in the network, the latter quantifies the time it takes to pass a signal to *every* node in the network. An analytical method is summarized that connects the degree distribution to the prediction of first passage times. This method is only valid for *locally tree-like* networks, i.e. networks

with vanishing clustering. Theoretical advancements concerning the efficient estimation of the cover time for arbitrary networks are an active field of research.

In Ch. 4, new small-world models are devised on the basis of the existing models. The models' control parameters are subsequently mapped to each other, showing that all follow similar structures and can be interpreted to be of modular hierarchical structure within certain regions of their control parameters. The discussion entails the generalization of the small-world modeling approach which includes all small-world models discussed. All models' structural control parameters regulate the trade-off between connecting nodes of short- and long-distance, varying a network's structure from locally connected to randomly drawn interactions. On this structural trajectory the clustering coefficient decreases for all network models. For three of the models the degree variance concurrently increases while it stays constant in a fourth model.

Based on these models, first passage time observables are investigated in Ch. 5. It is shown that on three of the network models, a minimum in the pair-averaged first passage time emerges in between the two structural limits. A new heuristic is found, which for the first time allows to estimate the pair-averaged first passage time based on both degree distribution and clustering coefficient. Based on this heuristic, the emerging minimum is explained by two countering effects: When manipulating a structure from being locally clustered to becoming more random, the decreasing clustering coefficient yields a decrease in the first passage time. At the same time, the increasing randomization introduces more nodes with a low number of connections which are harder to find randomly, therefore increasing the first passage time. The combination of the two effects explains the observed minimum. On the fourth small-world model, which has constant degree variance, the minimum in first-passage time does not emerge.

In Ch. 6, an optimal structure for fast distribution to *every* node is found for those three small-world models for which the pair-averaged first passage becomes minimal, similarly contradicting results from a corresponding average-medium approximation. The analysis is based on a newly-found heuristic which allows for the estimation of the cover time based on first passage time statistics and is demonstrated to be applicable to a variety of networks. On the fourth small-world model, the structural optimum vanishes because its degree distribution remains constant in between the structural limits.

Thus, combining the results of Ch. 5 and Ch. 6, one may not draw the conclusion that every modular hierarchical network is in a structural optimum for random messages to spread efficiently, negating the initially posed hypothesis.

Describing connections between humans on the basis of static networks as it is done in Part II of this work is an oversimplification of actual contact processes which are temporally resolved [35]. Such

simplifications are justified when processes taking their course in the system are associated with time-scales much longer than the time-scale of the system's structural change, for example for studying the spread of pandemics through global air traffic [3]. They become less useful when these corresponding time-scales approach values of similar order, as they might during the spread of an infectious disease [36]. For example, one might be interested in studying the spread of influenza on the static social network displayed in Fig. 1.1, assuming that every one of its links is a consequence of these two nodes meeting face-to-face at least once. This would, however, be a nonsensical endeavor: For such an infectious disease to spread from an infected to a healthy person, these two people have to be in spatial proximity for the several days during one of them is infected [37, 38]. Yet, many of the people linked in this static network do not meet face-to-face for years, rendering the static network a useless description of the human system in this context.

In order to obtain insight into the properties of actual human contact processes, two series of large-scale experiments measured face-to-face contacts between human individuals with a high temporal resolution in recent years [5, 39–42]. Their results show that people tend to spend time in small, densely connected groups, and change both their connectivity and activity based on a circadian rhythm. The size of these groups is heterogeneously distributed, and several temporal observables are following heterogeneous distributions, too, such as the duration of a face-to-face contact as well as the time a person stays isolated. This heterogeneity in interaction times is problematic because it defies the extraction of a definite time-scale which the time-scale of a disease can be compared to. Several approaches to model the aforementioned properties of real-world contact systems postulate that nodes spend time in groups and activate contacts based on heterogeneous memory-kernels, while ignoring the circadian rhythm [43–46]. This raises the question whether the heterogeneity of several network properties is indeed inherent to these systems, or whether they are emergent phenomena of underlying simple node behaviors that change with a circadian rhythm, as was for instance shown to be the case for heterogeneous inter-event times in e-mail communication [47, 48]. Furthermore, it is still unclear how to properly classify the structural and temporal properties of these networks over time and their influence on epidemic spreading. Thus, in Part III of this work I investigate in which way properties of temporal networks emerge as based on simple Poissonian processes with circadially varying rates and how this description influences the spread of infections.

First, temporal networks and the empirically collected contact data of references [5, 41, 42] are summarized in Sec. 2.3. Subsequently, existing modelling approaches are described in Sec. 2.4. In Ch. 7, a new temporal network model is formulated which is based on two fundamental node behaviors that are expressed with two rates: an active disconnection rate and an active reconnection rate. A variety of net-

work properties can be computed analytically. It is shown that the network fragmentation process to form densely connected groups emerges from the fundamental node behaviors. Subsequently, a new method is developed to infer the temporal variation of the dis- and reconnection rates from empirically collected data. It is found that all networks can be interpreted to follow similar trajectories in a newly-defined state space with one structural dimension and one temporal dimension. The structural dimension is quantified by the expected mean degree and the temporal dimension is reflected by the active reconnection rate per node, thus allowing for the first time to classify temporal networks as a point in this space where structural and temporal mixing are given by average node properties. Simulating the devised model with these varying rates, it is shown that the heterogeneity of the group-size distribution arises due to the circadian rhythm and a certain experimental practice. This practice entails the discretization of ime by binning contacts to equally-sized time intervals. The heterogeneity that temporal observables display in real-world data cannot be replicated, suggesting that other effects such as inherent node and link heterogeneity, or memory effects might be at play.

Concerning the spread of infectious diseases, much research has been devoted to study compartmental epidemiological models such as the susceptible-infected-susceptible (SIS) model on static networks (see [11, 49–52] and [9, pp. 192 ff.]). In this disease modeling approach, an individual can be in either of two states: infected or susceptible to infection. An infected node becomes susceptible with a recovery rate and a susceptible neighbor of an infected node becomes infected with an infection rate, which quantifies the infectiousness of a disease in terms of the speed with which it spreads within a system. Such compartmental models have been successfully used to investigate e.g. influenza spreading in a boarding school [53], the rise of measles infections [54, pp. 8 ff.], or how diseases spread through air traffic [3, 4]. In frameworks where nodes in a network correspond to single individuals, the widely applied individual-based Markov-chain approximation (MCA) yields insight into the epidemic response [49–51]. The epidemic response curve quantifies the number of infected people in equilibrium as a function of the *basic reproduction number*—this control parameter is proportional to the infection rate and reflects the infectiousness of a disease as the average number of secondary infections a 'patient zero' causes before they recover [55]. For many systems, the epidemic spreading process is subject to a phase transition: below a critical value of the basic reproduction number the system is effectively disease-free and resistant to outbreaks, above the critical value an endemic state is approached. This critical value is called the epidemic threshold and represents the susceptibility of a system to be subject to an epidemic [55]. Using the individual-based Markov-chain approximation, it has been shown that a heterogeneous degree distribution lowers the epidemic threshold such that even weak diseases can become endemic

in such systems, which was confirmed with simulation results using a discrete-time approximative SIS simulation algorithm instead of the original Poisson process [49, 51]. Skepticism towards the generality of this result, however, was raised, because the MCA assumes that infection states of neighboring nodes have zero covariance, an assumption which was shown to be violated in a number of systems [52]. Furthermore it was argued that the discrete-time approximative simulation method reduces the covariance between neighboring infection states such that the correspondence between simulation and MCA increases in these systems [52].

In the context of temporally resolved face-to-face contact networks, only few efforts have been undertaken to study the consequence of this temporal description on both the epidemic response and the epidemic threshold, as opposed to the description as a static, averaged network [45, 56, 57]. They found that in almost all empirical and model temporal networks, the epidemic threshold decreases when the time-scale of the disease (as quantified by the recovery rate) approaches the time-scale of the structural network change. This result is remarkable because it would mean that the shorter an individual is infected, the more likely the disease is to become endemic in a system—indeed a highly counter-intuitive result. These results are, however, all based on the MCA and *discrete-time* approximative simulations. Due to the objections described above, which were raised for the application of these methods on static networks [52], one may wonder about the validity of these results when the actual Poisson spreading process in continuous-time is considered on temporal networks. Furthermore, it is still unclear how the fragmented network structure influences the spreading process, especially when this structure is varying with a circadian rhythm. To find answers to these questions, Part III of this work concludes with a detailed analysis of spreading processes in temporally resolved contact networks.

To this end, modeling and simulation approaches to study the SIS model on static and temporally varying networks are introduced in Sec. 3.2, including a method to simulate the exact Poisson process in continuous time. In Ch. 8, this method is used to simulate the exact SIS process on temporally varying model networks in continuous time and to compare the results to predictions obtained using the MCA. In order to systematically study and compare the influence of structural and temporal properties on the spreading process, the networks are considered to be at a constant point in the state space newly defined in Ch. 7. For a purely random network model it is found that the MCA only approximates simulation results well in regions where a node is connected to many other nodes or where it often changes its environment. Reducing either the mean degree or the node mixing rate, simulation results show an increasing epidemic threshold, essentially making the system less susceptible to disease. In contrast, the MCA predicts a decreasing threshold which would mean the system becomes more susceptible, in direct conflict to the simulation results. An analysis of group-structured networks reveals

similarly conflicting results. Expanding the discussion to real-world temporal networks and their corresponding surrogates of Ch. 7 it is shown that the consideration of a circadian activity rhythm introduces complex behavior in the epidemic response curves. Since the networks can be interpreted to follow trajectories in the newly-defined state space they might enter regions where they are sparse and slowly changing enough for a disease to vanish. Thus, sudden jumps in the epidemic response curves may appear, an effect previously unobserved and non-replicable by the MCA.

The results of this work therefore suggest that the development of new theoretical approaches is necessary to accurately study the spread of diseases in temporally resolved contact networks.

Ultimately, the work presented here advances the understanding of *basic* dynamic processes in certain human systems modeled as networks. It shows how to connect network structure to the analysis of spreading processes, introduces several methods to analyze temporal and structural properties of those systems, and shows that some of the methods applied to understand basic processes are flawed. While the field is currently expanding to study more and more complex processes on increasingly complex structures, this work underlines the importance of taking a step back and understanding the basics before moving on and understanding the complex.

Part I

Human Systems as Networks: Theory & Data

2
Networks

As indicated in the introduction, networks have proven to be a powerful tool to analyze spreading processes in real-world systems. This first part of the dissertation therefore summarizes both concepts and open questions in current research regarding networks and spreading dynamics. This chapter focuses on the theory behind conceptualizing real-world situations on the basis of a network picture, leaving basic spreading processes for Ch. 3. The concepts and data introduced here present a comprehensive and in no way exhaustive overview, but rather focus on points which are of importance for the results presented in Part II and Part III.

Sec. 2.1 introduces several network definitions and their corresponding mathematical formulations, including a comprehensive description of networks with modular hierarchical structure in Sec. 2.1.7. These properties are the basis for the discussion of certain network models in Sec. 2.2 and in particular the small-world effect in Sec. 2.2.4. Subsequently, the discussion is extended to networks that vary over time in Sec. 2.3. The empirical data on which the discussion of Part III is based is presented in Sec. 2.3.3 and existing temporal network models are briefly discussed in Sec. 2.4.

2.1 *Network Properties*

Networks have certain structural properties that can be used to e.g. estimate the outcome of dynamic processes running on them, or to distinguish the heterogeneity of nodes in a network. Both of these will be relevant within the scope of this thesis. To this end, the adjacency matrix will be introduced in this section, based on which the node degree distribution and the global clustering coefficient will be defined. Furthermore, the average shortest path length and the concept of networks consisting of multiple components will be summarized. While other network properties and descriptions exist and are useful, their definitions are not of relevance for this work and hence the discussion will be constrained to those mentioned before.

2.1.1 *The Adjacency Matrix and Some Other Network Definitions*

A network is a representation of a set of entities and a set of connections between any *two* of those entities.[1] The entities will be called

[1] The distinction of a connection of two entities has to be made since there are descriptions where contacts of three or more entities are considered called 'hypergraphs', however, these are not of concern in this work.

"nodes" in the following and the connections will be referred to as "links", "edges", or "contacts".[2] When a node u is connected to node v by a link, v will be called a "neighbor" of u's and vice versa. All nodes which are connected to node u by a link will form u's *neighbor set* $\mathrm{Nei}(u)$, sometimes also called *neighborhood*.

[2] However "contact" will usually be used in the context of temporal networks, see Sec. 2.3.

Since links only ever connect two nodes at once, one can describe a network by a so-called *adjacency matrix*, which is used to keep track of the connections and non-connections between nodes and will prove to be a useful mathematical object later on. In a network of N nodes, the adjacency matrix $\boldsymbol{A}$ is an $(N \times N)$-sized matrix. In the simplest case, a network is unweighted an undirected. Then, the adjacency matrix is given as

$$A_{vu} = \begin{cases} 1 & \text{if } v \text{ is connected to } u, \\ 0 & \text{otherwise.} \end{cases}$$

In an undirected network, the existence of an edge (u, v) means that the edge points from node u to node v, and from node v to node u. Hence, the adjacency matrix is symmetric as $\boldsymbol{A} = \boldsymbol{A}^T$. An example is displayed in the upper left panel of Fig. 2.1. The total number of edges is given by counting all one-entries of A as

$$m = \frac{1}{2} \sum_{u=1}^{N} \sum_{v=1}^{N} A_{vu}. \tag{2.1}$$

Note that while the symbol m will be mainly used to refer to edge counts, the symbol e will be used in Ch. 7.

All empirical and synthetic networks discussed in this thesis will be undirected and most will be unweighted. Exceptions are random walk transition matrices which can be interpreted to be weighted and directed, which is why the discussion is briefly extended to directed and weighted networks below. For weighted networks and each pair of nodes (u, v), the entry A_{vu} of the adjacency matrix describes the strength of the interaction[3] between node v and node u. Usually for an entry $A_{vu} = 0$ it is implied that there is no connection between u and v. In contrast, if $A_{vu} > 0$, a connection exists[4] and its strength is A_{vu}. An example for this kind of adjacency matrix is shown in the lower left panel of Fig. 2.1.

[3] What "strength of the interaction" actually means heavily depends on the context in which a network description is chosen to be applied.

[4] Note that the limitation of A_{vu} being positive is relatively arbitrary. It is made here since for mosts contexts in this thesis, some kind of spreading probability will be proportional to this weight. In order to properly define and normalize this probability, it is necessary that this weight is non-negative. Other networks which are based on e.g. correlations or reaction rates might very well contain negative edge weights.

In a directed, weighted network, for each pair of nodes (u, v) the entry of the adjacency matrix A_{vu} describes the strength of the interaction on node v given that the source of the interaction is node u. An entry $A_{vu} = 0$ usually implies u cannot influence v or that a commodity cannot spread from node u to node v. In contrast, if $A_{uv} > 0$, node v can be influenced by node u. Hence, the adjacency matrices are not necessarily symmetric such that $\boldsymbol{A} \neq \boldsymbol{A}^T$. An example is shown in the lower right panel of Fig. 2.1.

Empirical networks are often both directed and weighted. In order to analyze the contact structure only, those are often converted to undirected and unweighted networks. This is useful when, for instance, new analysis methods are developed which are usually sim-

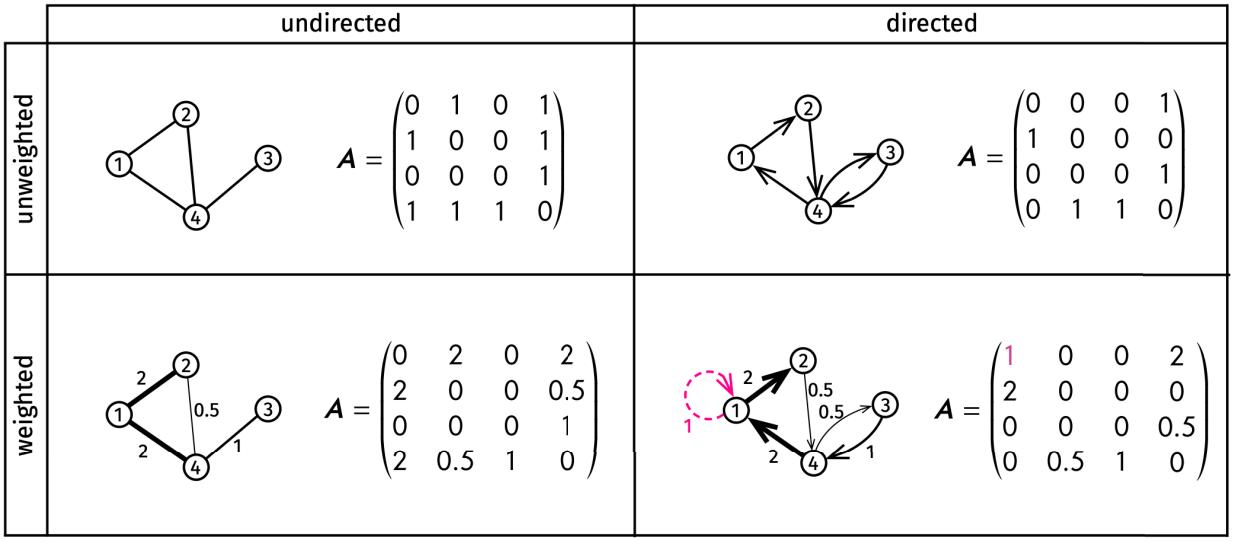


Figure 2.1: Example networks from the space {undirected, directed} × {unweighted, weighted} and their respective adjacency matrices. Additionally marked in pink is a self-loop which might occur.

pler to evaluate on the latter. Using the Heaviside step function

$$\Theta_t(x) = \begin{cases} 0 & x \leq t, \\ 1 & x > t, \end{cases}$$

a weighted directed network defined by the adjacency matrix $\boldsymbol{A}^{\mathrm{WD}}$ can be converted to an unweighted directed network $\boldsymbol{A}^{\mathrm{UD}}$ as

$$A^{\mathrm{UD}}_{vu} = \Theta_t(A^{\mathrm{WD}}_{vu}),$$

also called a *thresholding procedure* because all weights above a threshold t are converted to unweighted edges and all edge weights below or equal to the threshold will be deleted. Further simplifying the structure, the network can be made undirected by applying the step function once again to obtain

$$A^{\mathrm{UU}}_{vu} = \Theta_0(A^{\mathrm{UD}}_{vu} + A^{\mathrm{UD}}_{uv}).$$

Then, the matrix $\boldsymbol{A}^{\mathrm{UU}}$ is symmetric and its entries consist of ones and zeros only. This procedure will be useful in Ch. 6 where empirical networks will be reduced to their undirected and unweighted counterparts in order to test the newly developed method for computing the mean cover time.

2.1.2 *Node Degree and Degree Distribution*

Nodes can play different roles depending on their properties and the situation one wishes to analyze. One of a node's simple properties is its total number of neighbors, a number which will be referred to as its *degree* in the following. The degree conveys some important information about a node's role for certain dynamic processes. Consider, for example, a random search process in which one aims at locating a target node by randomly jumping from node to node following

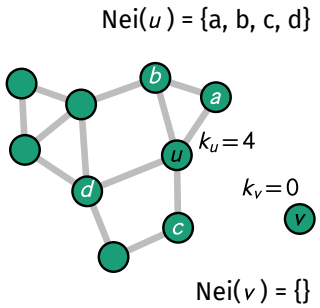


Figure 2.2: A small network and some node neighborhoods. The node u has the neighbor set $\mathrm{Nei}(u) = \{a, b, c, d\}$ and is therefore associated with the degree $k_u = |\mathrm{Nei}(u)| = 4$, while the node v is not connected to any other node and therefore has degree $k_v = 0$.

connections between them. In such a process, a node of large degree will ususally get found in a smaller number of steps than a node with degree one, simply because the latter it is only reachable by a single edge—and a node with degree zero will not be found at all. Situations like these will be analyzed in Sec. 3.1a and Part II of this work. Similarly, an infected node connected to a single healthy node is less likely to cause a single infection per unit time than an infected node connected to a hundred healthy nodes, for the whole reason of having a hundred-fold more possibilities to do so. Dynamic processes of this kind will be discussed in Sec. 3.2 and Ch. 8.

Based on the adjacency matrix $\boldsymbol{A}$ of an undirected network as defined above, the degree of a node u is defined as

$$k_u = \sum_{v=1}^{N} A_{vu}. \tag{2.2}$$

A node's degree is equally given by the cardinality of its neighborhood

$$k_u = \left|\mathrm{Nei}(u)\right|.$$

An example for neighborhoods and node degrees is shown in Fig. 2.2. In certain cases which will be further discussed later on, it suffices to analyze a network by assuming every node is approximately equal to an average node with *mean degree*

$$\langle k \rangle = \frac{1}{N} \sum_{u=1}^{N} k_u.$$

Using Eq. (2.2) and Eq. (2.1) one further finds

$$\langle k \rangle = \frac{2m}{N}$$

which means that the mean degree is also quantifying the total number m of edges in a network of N nodes. The mean degree further measures the density of a network. If $\langle k \rangle \ll N$, a network is called *sparse*. Increasing the mean degree corresponds to making the network *denser*. A network where at $N \to \infty$, the ratio $\langle k \rangle / N > 0$ is called *dense*.

One way to measure node heterogeneity is to compute the *degree variance*

$$\mathrm{Var}\,[k] = \frac{1}{N} \sum_{u=1}^{N} [k_u - \langle k \rangle]^2 = \left\langle k^2 \right\rangle - \langle k \rangle^2.$$

Typically, one speaks of a network with homogeneous node degree when the degree variance is of a similar order as or lower than the mean degree. A more reliable way to investigate node degree heterogeneity is to analyze the whole *degree distribution*. The degree distribution is denoted by P_k which represents the probability that a node picked at random has degree k.[5] In network models, the degree distribution may be computed to follow a given functional form, but for finite-size samples or empirically connected networks, P_k usually refers to a normalized histogram.

[5] Technically, the distribution is a probability mass function (pmf) because the domain is discrete, but this will not be of concern here.

The example network shown in Fig. 2.2 has a degree distribution which is depicted as a histogram in Fig. 2.3a. The single node v with degree $k = 0$ is associated with $P_k = 1/10$. There are no nodes with only a single connection.

The author's Facebook friend network shown in Fig. 2.3 follows an exponential distribution with mean $\langle k \rangle = 10.98$ in Fig. 2.3b, which will be expressed as $k \sim \mathcal{E}(1/\langle k \rangle)$ in the following. The exponential distribution follows the probability density function $P(k) = \langle k \rangle^{-1} \exp(-k/\langle k \rangle)$ with degree variance $\mathrm{Var}[k] = \langle k \rangle^2$, which fits well to the Facebook network's empirical distribution with degree variance $\mathrm{Var}[k] = (10.83)^2$. As the variance is one magnitude greater than the mean degree, the network is considered to have a heterogeneous degree distribution. Particularly, a few nodes have rather large degrees of $k = 62$ or $k = 63$. These will be referred to as *hubs.*

2.1.3 *A Word on Heavy-Tailed Distributions*

At this point it seems useful to introduce the notion of homogeneous, heterogeneous, and heavy-tailed distributions in the way they are used here. In certain contexts, homogeneous distributions refer to situations where the distribution's variance is of the same order as its mean, as it was used in the context of degree distributions. In Part II, network models are introduced where the node degree variance increases from $\mathrm{Var}[k] = 0$ to $\mathrm{Var}[k] = \langle k \rangle$. In this case, it will be said that the degree distribution becomes *more* heterogeneous even though the limit $\mathrm{Var}[k] = \langle k \rangle$ is considered homogeneous, the reason being that certain dynamics react sensitive even to small changes in the degree distribution. The exponential distribution, which will be used many times in this work, has variance $\mathrm{Var}[x] = \langle x \rangle^2$. For degree distributions, this would already be considered heterogeneous because the outcome of spreading and random walks is usually compared to homogeneous structures (see e.g. Sec. 3.2).

However, in the case of waiting time distributions the exponential distribution $\tau \sim \mathcal{E}(\omega)$ will be considered to be homogeneous, because it decays more quickly than so-called heavy-tailed distributions, such that the waiting time between events is associated with a single time scale ω^{-1}. As an illustration, consider a waiting time distribution comprised of multiple time scales distributed as $\omega \sim \mathcal{E}(\omega_0^{-1})$ where $\omega_0^{-1} = \tau_0$. In this case, the waiting time distribution is given as

$$P(\tau) = \int_0^\infty d\omega\, \omega \exp(-\omega\tau) \frac{1}{\omega_0} \exp(-\omega/\omega_0)$$
$$= \frac{\tau_0}{(\tau + \tau_0)^2}.$$

This resulting distribution is part of a family called *heavy-tailed* distributions which means that a slowly decaying tail prevents the existence of certain moments which diverge. Values distributed according to heavy-tailed distributions can therefore not be associated

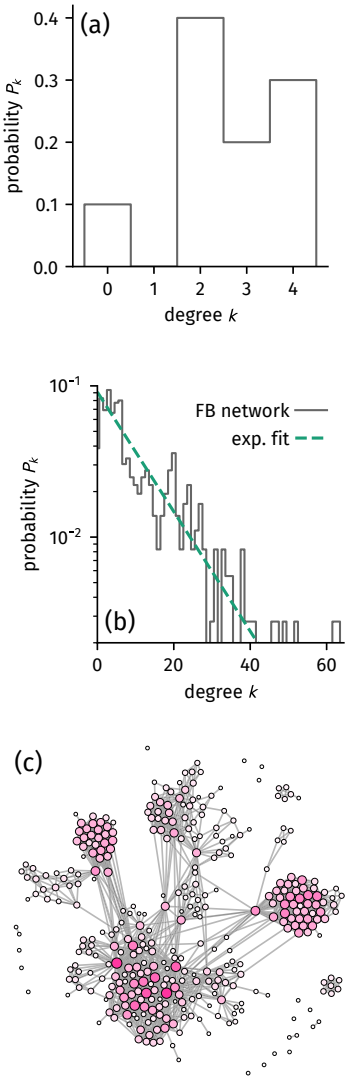


Figure 2.3: Degree distributions of (a) the exemplary network shown in Fig. 2.2 and (b) the Facebook network shown in Fig. 1.1 and (c). **(a)** In the example network, the probability to sample a node of degree $k = 0$ is $P_k = 1/10$ because a single node of the $N = 10$ nodes comprising the network has zero degree. **(b)** The degree distribution of the author's Facebook friend network may be modeled with an exponential distribution. A maximum likelihood fit reveals an exponential decay with mean $\langle k \rangle = 10.98$. The network's degree variance $\mathrm{Var}[k] = (10.83)^2$ fits well with the exponential hypothesis which predicts $\mathrm{Var}[k] = \langle k \rangle^2$. **(c)** The Facebook friend network with each node colored according to its degree with $k = 0$ corresponding to white and $k = 63$ corresponding to pink.

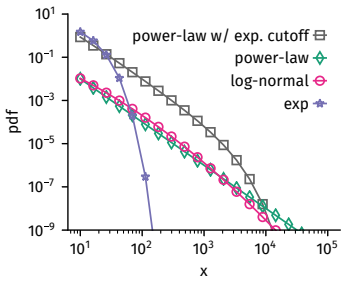


Figure 2.4: Different model distributions to illustrate heavy tails. Shown here are an exponential distribution (not heavy-tailed), and power-law, log-normal, and power-law with exponential cutoff (heavy-tailed distributions).

with a typical scale, which is a problem in a number of contexts. For instance, the mass of an adult human being follows a homogeneous distribution, associated with a typical scale: the average mass, e.g. $\approx 70\,\mathrm{kg}$ for German, Austrian, and Norwegian conscripts [58]. For an application such as the transport of 100 soldiers to another base, it therefore suffices to compute the total amount of mass moved as $100 \times 70\,\mathrm{kg}$. On the other hand, stock price returns follow a heavy-tailed distribution [59]. Therefore, if a stock broker places 10 bets on stocks and predicts their expected return as the 10-fold of the average return, it is very likely that they go bankrupt on highly probable low returns before they hit a high return. In this case, the mean is not representative of the distribution, which lacks a typical scale.

While the definition for heavy-tailed distributions given above is straight-forward in theoretical contexts, empirically collected data is usually of finite size and as such, all moments are always computable. Therefore, in this thesis, distributions will be called *heavy-tailed* when they roughly follow a power-law $P(x) \propto x^{-\mu}$ in a dominant region of the domain (see Fig. 2.4). Famous examples include the log-normal distribution or the power-law itself. Power-laws with a cutoff-behavior effectively limit the domain to an upper bound, however will be included in the heavy-tailed class here because while there is an upper bound, it is still not possible to associate the distribution with a single scale.

2.1.4 *Clustering Coefficient*

Social networks are networks in which nodes represent individuals of a social species and links represent relationships between them, e.g. friendships, frequent grooming, or just an increased amount of time spent together. In those networks, it is not unusual to find high occurrences of triadic closure—meaning that if an individual is 'friends' with two other individuals, it is likely that those two individuals are 'friends', as well. One way of quantifying this is to measure the conditional probability that any two nodes are connected by a link given that they are both connected to a focal node. In an undirected unweighted network with adjacency matrix $\boldsymbol{A}$, this clustering coefficient of a single node i is defined as

$$C_i = P\left[A_{iu}A_{uv}A_{iv} = 1 | A_{iu}A_{iv} = 1\right].$$

Since the conditional probability that two nodes u, v are connected to the focal node i given that they are part of the corresponding triangle is always $P\left[A_{iu}A_{iv} = 1 | A_{iu}A_{uv}A_{iv} = 1\right] = 1$, this clustering coefficient evaluates to

$$C_i = \frac{P[A_{iu}A_{uv}A_{iv} = 1]}{P[A_{iu}A_{iv} = 1]}.$$

Any two nodes which are not the focal node may connect to the focal node to form a so-called *two-star* of form •–•–• where the focal node is the central node. The total expected number of two-stars

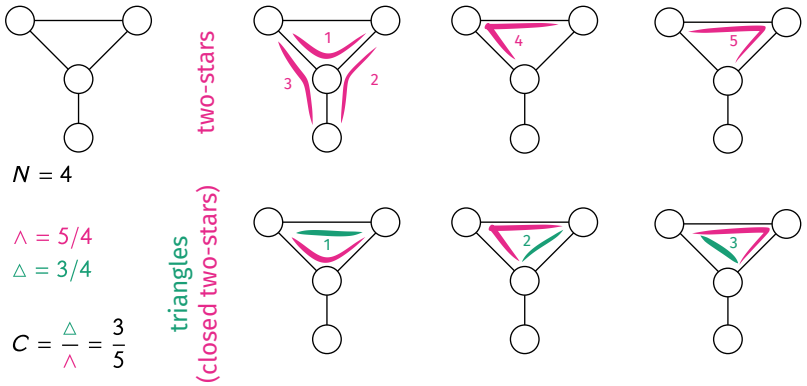


Figure 2.5: An illustration of the clustering coeffcient. In the depicted network of $N = 5$ nodes, the number of two-stars per node is $\wedge = 5/4$ and the number of triangles per node is $\triangle = 3/4$. Consequently, the global clustering coefficient is $C = 3/5$. Note that the global clustering coefficient may differ from the average clustering coefficient per node which here is $\langle C_i \rangle = (1/3)(1/3+1+1) = 7/9$.

with focal node i in the center is therefore given as

$$\wedge_i = \frac{1}{2}(N-1)(N-2) \times P[A_{iu}A_{iv} = 1]. \tag{2.3}$$

Similarly, the total expected number of triangles that the focal node is part of is given by the amount of two-stars which are triadically closed

$$\triangle_i = \frac{1}{2}(N-1)(N-2) \times P[A_{iu}A_{uv}A_{iv} = 1]. \tag{2.4}$$

With these definitions, the clustering coefficient of focal node i can be computed as the ratio of expected closed two-stars

$$C_i = \frac{\triangle_i}{\wedge_i}.$$

An illustration of those observables is shown in Fig. 2.5. In certain network models, each node i is statistically equivalent, meaning that their average properties will be equal and thus independent of the node. In this case, $\wedge_i = \wedge_j \equiv \wedge$ as well as $\triangle_i = \triangle_j \equiv \triangle$ such that

$$C_i \equiv C = \frac{\triangle}{\wedge}.$$

This is the definition of the clustering coefficient which will be used to measure triadic closure in the network models defined in Ch. 4. For empirical networks or samples of a network, the clustering coefficient will be defined as the global probability of triadic closure given by the ratio of closed two-stars such that again $C = \triangle/\wedge$ where

$$\triangle = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\sum_{u=1}^{N}\sum_{v=1,v\neq u}^{N} A_{iu}A_{uv}A_{iv}$$

$$\wedge = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\sum_{u=1}^{N}\sum_{v=1,v\neq u}^{N} A_{iu}A_{iv}.$$

Note that in this case however, the global clustering coefficient will *not* be equal to the mean clustering coefficient of a focal node $\langle C_i \rangle$ because in general $\langle \wedge/\triangle \rangle \neq \langle \wedge \rangle / \langle \triangle \rangle$. One further notes that a single node of degree k will be the center of $(1/2)k(k-1)$ two-stars and as such

$$\wedge = \frac{1}{2}\left[\left\langle k^2 \right\rangle - \langle k \rangle\right] = \frac{1}{2}\left[\mathrm{Var}\left[k\right] - \langle k \rangle + \langle k \rangle^2\right].$$

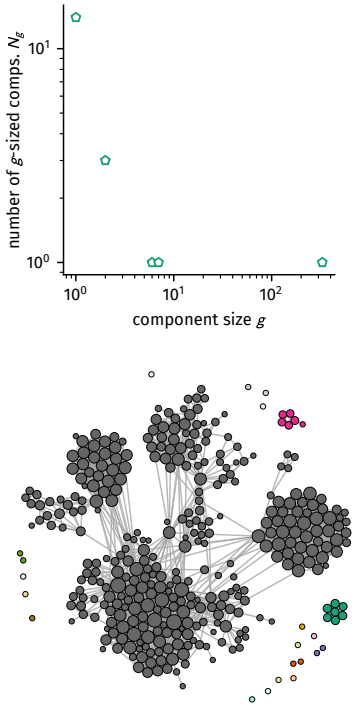


Figure 2.6: **(Top)** component size distribution of the author's Facebook friend network with values $n_1 = 14$, $n_2 = 3$, $n_6 = n_7 = n_{329} = 1$, and $n_{g'} = 0$ for $g' \notin \{1, 2, 6, 7, 9\}$. Eq. (2.5) yields the correct number of nodes in the network $N = (1 \times 14) + 2 \times 3 + 6 \times 1 + 7 \times 1 + 329 \times 1 = 362$. **(Bottom)** the corresponding network, where each node is colored according to the component it is part of.

In this work the clustering coefficient will be used to quantify how useful certain network models are to model social networks when the only criterion for "usefulness" is non-zero probability of triadic closure.

The author's Facebook friend network has $\wedge = 113.5$ two-stars per node, $\triangle = 58.26$ triangles per node, and is therefore associated with a global clustering coefficient of $C = 0.51$.

2.1.5 *Networks Composed of Multiple Components*

In some undirected networks, not every target node can be reached from every source node. If this is the case, the network consists of multiple disconnected components with drastic consequences for certain dynamics: For instance, since some nodes are not reachable, the average mean time to reach every node diverges. In order to avoid such difficulties, one typically extracts the *largest component* of a network. In order to find all components of a network, the following algorithm can be used:

1. Mark each node as not being part of a component and set the number of components to zero.
2. Iterate through all nodes
   (a) For each node, check whether it is marked as being part of a component.
   (b) If it is, proceed with the next node in the list of all nodes.
   (c) If it is not, mark the node with the next component number and increase the number of components by one. Then:
      i. For each of the node's neighbors, check whether it is marked as being part of a component.
      ii. If it is, proceed with the next neighbor.
      iii. If it is not, mark the neighbor with the current component number, set the focus to this neighbor and proceed with steps (i)-(iii).

This recursive algorithm yields a set of distinct components, each component being a set of nodes. One may quantify a network's fragmentation using the group-size distribution n_g which refers to the number of groups of size g. In any case, the union of all components contains all of the network's nodes. Therefore, the group-size distribution is constrained by

$$N = \sum_{g=1}^{N} g \times n_g. \tag{2.5}$$

In connection to bond percolation [32, pp. 347 ff.], above a critical number of edges in the system (or a critical number of average connections per node, which is equivalent), a network has a largest component of size $g_{\max} \sim \mathcal{O}(N)$. In Part II, this component will be

extracted to avoid divergences in temporal observables. For the author's Facebook network shown in Fig. 1.1, the component analysis shows that there exists a single largest component of size $g_{\max} = 329$ and several significantly smaller components (see Fig. 2.6). A brief discussion of the critical mean degree for a largest component to emerge will be held in Sec. 2.2.2.

In the context of temporal face-to-face networks, a single snapshot often consists of a multitude of small-sized densely connected components (see Sec. 2.3.3). A major contribution of this thesis will be to investigate the influence of this group-size distribution on the spread of diseases in Ch. 8.

2.1.6 *Shortest Path Lengths*

For several problems approachable using network theory, it is important to know how "far" two nodes are away from each other. Some examples include the navigation of public transport, the minimal arrival time of data packets sent over the internet, or the risk of disease import. Usually, the notion of distance between two nodes is defined as the minimal number of edge traversals needed to travel from a source node u to a target node v—the *shortest path length* s_{uv}. This shortest path length can be found using Dijkstra's algorithm [60].[6] Considering search algorithms, where starting at a source node u, the task is to find a target node v, the shortest path between those nodes corresponds to the optimal search time. If a search algorithm has maximum information about the network, it can apply Dijkstra's algorithm to find the shortest path and simply follow it. Usually, however, search processes will only have access to local information or, in the worst case, no information at all—the search process will be random. Computing the average number of steps it takes a random search to arrive at a destination therefore places an upper bound on search times of any informed search.

In order to quantify the general "closeness" of any two nodes in a network, it is common to compute the *average shortest path length*

$$\langle s \rangle = \frac{1}{N(N-1)} \sum_{u=1}^{N} \sum_{v=1}^{N} s_{uv},$$

where $s_{uu} = 0$. Networks where the average shortest path length scales as $\langle s \rangle \propto \log N$ with the number of nodes N are called *small-world networks* [1] because compared to e.g a lattice, only a small number of steps are necessary to travel from one node to another (on average). The small-world effect will be discussed more detailed in Sec. 2.2.4 and Part II of this work where the *pair-averaged first passage time* will be computed as an upper bound for the average time of an informed search, thus generalizing the small-world effect. The pair-averaged first passage time will be introduced in Sec. 3.1.5.

[6] Because shortest paths are only of minor concern in this thesis the explanation of this algorithm will be omitted.

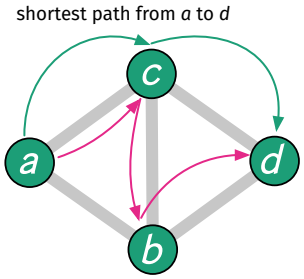


Figure 2.7: An illustration of a shortest path. There is an infinite number of paths from node a to node d. Highlighted are the paths $a \to c \to d$ and $a \to c \to b \to d$. The shortest path length from a to d is $s_{da} = 2$. There are $N(N-1) = 12$ target-source pairs in this network. Choosing nodes b and c, both contribute three shortest paths of length $s = 1$ because they are connected to all other nodes. The nodes a and d both miss a connection to a single other node and therefore contribute two shortest paths of length $s = 1$ and a single shortest path of length $s = 2$ to the average. Hence, the average shortest path length of this network is $\langle s \rangle = (2 \times (3 \times 1) + 2 \times (2 \times 1 + 1 \times 2))/12 = 7/6$.

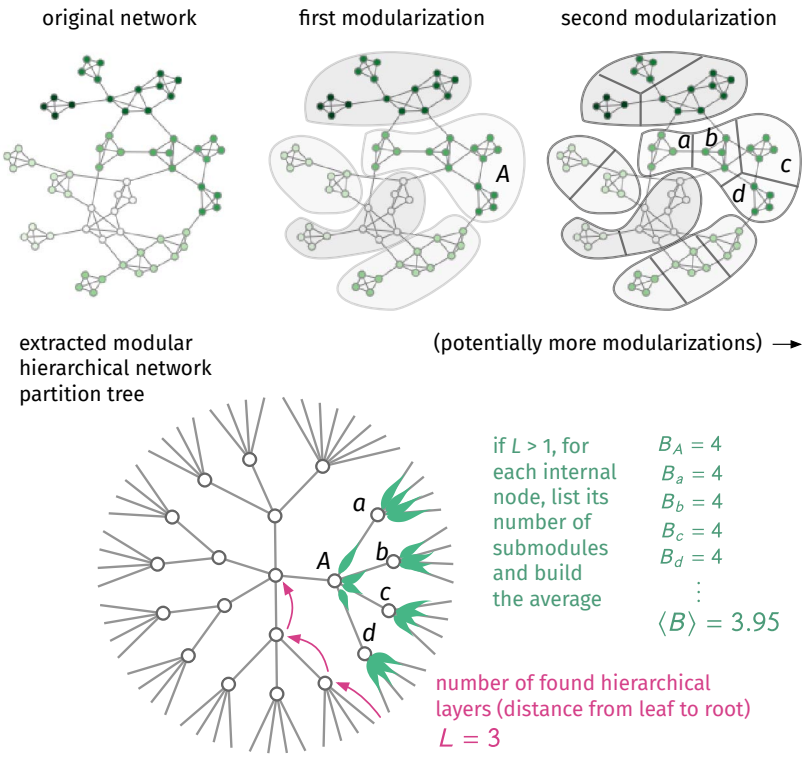


Figure 2.8: Illustration of how networks can be viewed as a system of hierarchically organized modules and how these structures will be analyzed in this thesis. **(Top row)** A network is first partitioned into regions which are "closer" to each other. In a second step, each of these regions is again partitioned into regions which are "closer". Further partitions within those regions are potentially possible. In this thesis, the hierarchically nested version of *Infomap* is run to find a modular hierarchical network partition, where nodes being "close" refers to a random walker spending a substantial amount of time in their region before transferring to some other parts of the network (see App. A.). **(Bottom row)** this network partition can be displayed as a tree of height L. The central node which does not have any predecessors is called *root* and contains all subregions of the total network. An *internal node* is any partition which contains regions or nodes of the original network. A node without any successors is called a *leaf* and refers to a node in the original network. The number L of found hierarchical layers is the maximum number of steps it takes to traverse from any leaf to the root while only going upwards. The height L of the found tree is one of the measured structural observables to quantify modular hierarchical structure. Furthermore, each internal node I (or submodule) carries B_I submodules. From these, the mean number $\langle B \rangle$ of submodules per internal node is computed (including its standard deviation in later parts of this thesis).

2.1.7 *Modular Hierarchical Structure*

A ubiquitous feature of real-world networks is that they display community structure, meaning that one may partition a network into a collection of modules, each representing a distinct set of nodes. Nodes in a module are considered to be "closer" to each other than to nodes in other modules, often quantified as having more connections to nodes within the module than to nodes in other modules [61–64]. Various studies have further considered the presence of hierarchies of these communities, describing networks that contain further submodules within modules [14, 24, 28, 29, 61], as illustrated in Fig. 2.8. A remarkable number of natural and social systems can be described as a hierarchy of modules [28, 65], such as brain networks [15–20], metabolic and cell networks [14, 22, 23], human transport networks [24, 25], ecological systems [14, 26, 27] and social networks [19, 65] (examples given below). When analyzing empirical networks for their modular hierarchical structure, one typically refers to the found partition by means of an inferred tree (see Fig. 2.8), where the *root* refers to the zero-level partition which contains all nodes, an *internal node* refers to any module containing submodules or nodes, and a *leaf* of the tree corresponds to the actual nodes in the analyzed network. The tree's maximum distance L between any leaf and the root is called its *height* and is synonymous with the number of *hierarchical layers*. Each internal node I (corresponding to a module) possesses a number B_I of submodules.

A significant amount of research has focused on developing al-

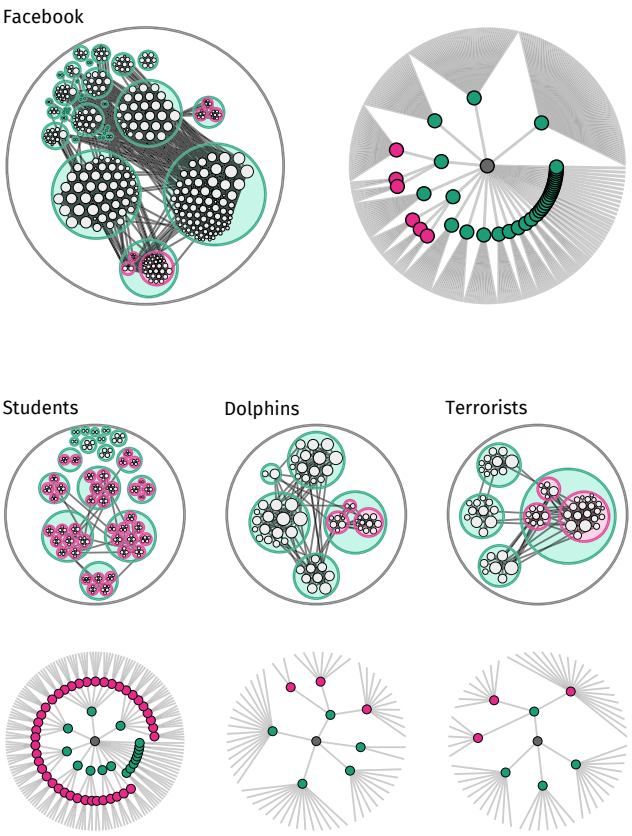


Figure 2.9: Modular hierarchical structure of various social networks, where nodes (white disks) refer to individuals and a link between two of those individuals existed when they spent a certain amount of time with each other or exchanged information. Colored circles refer to modules. A colored circle within a circle refers to a submodule within a module. More detailed information about the networks and the inferred modular hierarchical partitions can be found in Tab. 2.1.

gorithms to partition empirical networks in a modular hierarchical way [14, 19, 24, 28, 29, 66]. However, the purpose of this thesis will be to investigate the question of why such hierarchical organization emerges and persists in the first place. Strategies to address this question involve the search for structural optima regarding the efficacy of certain dynamics, which will be further discussed in Sec. 2.2.4 and Sec. 3.1. While the focus of this work will therefore lie on the dynamical implications of modular hierarchical structure, it will sometimes be necessary to infer such structure. Since here, only qualitative descriptions are necessary, the particular choice of the inference algorithm is not of importance. Hence, a single algorithm called *Infomap* will be used [29, 67, 68], due to its high efficacy compared to other methods and its close relatedness to spreading dynamics. The algorithm infers structural organization by hierarchically grouping regions of nodes in which a random walker spends a considerable amount of time before entering other regions (random walks will be explained in detail in Sec. 3.1). *Infomap* minimizes the minimum description length of an ergodic random walk, which is given as an average number of bits and can be interpreted as a measure of how "compressible" the network is, given the found modular hierarchical network partition. A more detailed description of the algorithm's functionality is given in App. A.

Network	N	$\langle k \rangle$	C	$\langle s \rangle$	L	$\langle B \rangle$
B. F. Maier's Facebook friends network [6]	362	11.0	0.51	3.6	3	9.6
Cooperation between students [69]	185	3.4	0.53	6.5	3	4.1
Social interaction in dolphins [12]	62	5.1	0.31	3.4	3	7.8
Associations between terrorists [70]	62	4.9	0.36	2.9	3	8.6
Power grid in Western U.S. [1]	4941	2.7	0.10	19.0	7	5.5

Table 2.1: Network properties of a collection of real-world networks and relatedness to modular hierarchical structure. The Facebook network [6] refers to the network shown and discussed in Fig. 1.1. In the student network, nodes represent college students and a link is drawn when they worked together on any assignment [69]. The dolphin network consists of nodes representing individual dolphins and links represent relationships. A relationship was defined as dolphins spending a significant amount of time with each other during a 7 year observation period [12]. After the terror attacks against the U. S. on September 11, 2001, an association network between terrorists was drawn [70]. Each node represents a terrorist, and a link is a qualitative assessment of whether these terrorists where in any way in contact with each other. In the power grid network, a link between two nodes represents a high-voltage power line between two generators or transformer stations, represented by nodes [1]. Listed are the number N of nodes, the mean degree $\langle k \rangle$, the global clustering coefficient C, and the average shortest path length on the largest component $\langle s \rangle$; furthermore the evaluated number L of hierarchical layers and the mean number $\langle B \rangle$ of submodules per module, using the procedure described in the main text.

Regarding the modular hierarchical structure of real-world networks, examples are analyzed below, with a description of the networks and an overview of the results regarding the network structure given in Tab. 2.1. Four social networks of varying size are analyzed and shown in Fig. 2.9. In each of those networks, nodes represent individual people or animals and links represent any kind of association between them. All networks are found to be partionable by trees of height $L = 3$ and with average numbers of submodules per module ranging from $\langle B \rangle = 4$ to $\langle B \rangle = 10$. While being relatively sparse with a low mean degree, they have non-vanishing clustering and a relatively low average shortest path length. An example for a larger network of significant modular hierarchical structure is shown in Fig. 2.9, where the power grid of the Western U.S. is analyzed (description given in Tab. 2.1). Infomap finds a tree of height $L = 7$ with an average number of $\langle B \rangle = 5.5$ submodules per submodule. Clustering is lower than in the social networks and the average shortest path length between nodes is higher, even when considering the larger number of nodes.

Studying modular hierarchical structures, it is often found that modules might overlap, obfuscating the assignment of a node to a particular module. This can be resolved by studying modular hierarchical clustering of links rather than nodes [71]. These kind of overlapping structures evolve from nodes taking part in different kinds of social contexts such that links build modules according to the contexts they are activated in. While this is inarguably a more intuitive picture, the structure of many networks is still well-captured by node-centric approaches as was demonstrated here. Since the purpose of this work is to qualitatively investigate the influence of modular hierarchical clustering on dynamics, this node-centric approach will sufffice.

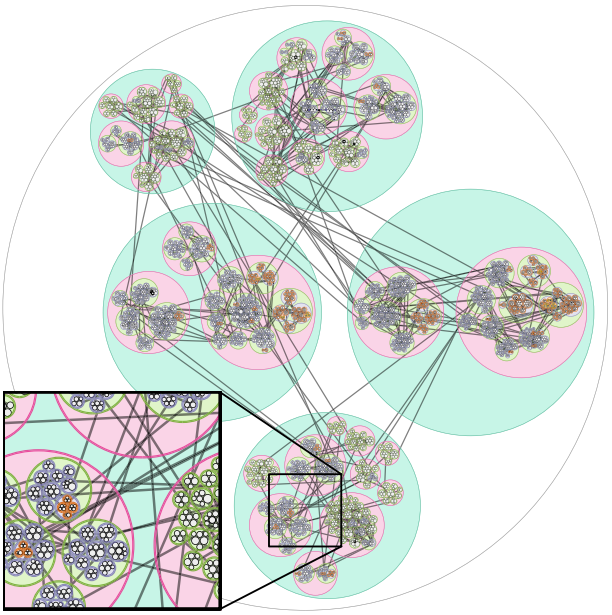


Figure 2.10: Modular hierarchical structure of the Western U. S. power grid where each node (white disk) represents a generator or transformer station. A link represents a high-voltage power line between a pair of those. Each colored circle corresponds to a module and each circle within in a circle is a submodule of the module represented by the enclosing circle. For this network, *Infomap* finds a modular hierarchical partition of height $L = 7$ with an average number of $\langle B \rangle = 5.5$ submodules per submodule. **(Inset)** usually, an average node of a submodule will have more connections to another node within the same submodule than to a node within the same module (but different submodule).

2.2 *Static Network Models*

There exist many network models, each to highlight different phenomena observed in real systems or to be used as reference models to compare effects on other networks. In this section, all network models used in this thesis will be introduced and their relevant properties and effects will be explained.

2.2.1 *Regular Networks and Lattices*

Very popular networks in physics and chemistry are any kind of lattice: nodes are positioned in symmetric ways on a geometric surface of integer-valued dimension and only connect to 'near' neighbors in a periodic manner.[7] An example is the two-dimensional square lattice of $N = a^2$ nodes with $a \in \mathbb{N}^+$, shown in Fig. 2.11 for $a = 8$. Lattices like these are popular because they represent a discretized version of space on which certain analyses are simplified or computationally more feasible. The most important feature concerning dynamics is that the average shortest path length is rather large on such networks: In order to traverse from one end of the network to another, one has to pass from nearest neighbor to nearest neighbor; there are no shortcuts. However, networks like the one shown in Fig. 2.11 do not possess clustering—the number of closed two-stars is zero. They are therefore less useful to study social networks in which $C > 0$.

A lattice model which solves this problem is the k-regular nearest-

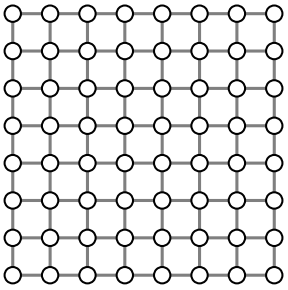


Figure 2.11: A two-dimensional square lattice of finite side length $a = 8$.

[7] There exist non-periodic lattices but traditional physicists are not very fond of those.

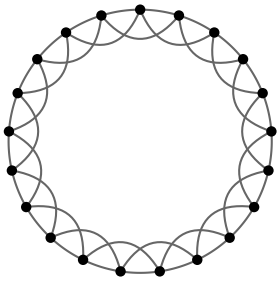


Figure 2.12: A k-regular nearest-neighbor lattice with $k = 4$.

neighbor lattice. For this model, nodes are positioned equidistantly on a ring and connect to their k nearest neighbors where $k = 2k'$ with $k' \in \mathbb{N}^+$ (see Fig. 2.12). In this network model, every node is connected to k nodes such that the degree distribution is $P_{k'} = \delta_{kk'}$. Each node is the center of $\wedge = (1/2)k(k-1)$ two-stars. In order to find the number of closures between those two-stars, note that links exist between all k of a focal node's neighbors which are of distance $< k/2$ on the ring, hence the number of triangles per node

2.13. Consequently, the clustering

$$\frac{(k-2)}{(k-1)}. \tag{2.6}$$

of $k = 4$ neighbors per node, the 2, similar to the value of the au- ($C = 0.51$). Similarly to the two- erage shortest path length is rela- de, a target node is between $d = 1$ e ring, hence the average shortest 4 on a ring. Introducing connec- , a walker may skip $k/2 - 1$ nodes k). The largest component of the ortest path length of $\langle s \rangle = 3.6$ and mean degree of $\langle k \rangle = 11.9$. While a plausible value of clustering for eproduce the network's short path e of $\langle s \rangle \approx 14$.

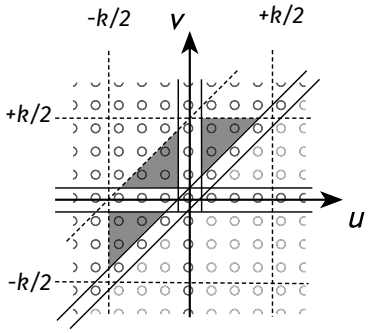


Figure 2.13: Evaluating the triangles per node in the k-regular nearest neighbor lattice. Consider a focal node at node position $i = 0$. Then, every node $-k/2 \leq u < 0$ and $0 < u \leq k/2$ is connected to node i. Consequently, every node $v \neq 0$ with $u + 1 \leq v \leq \min(k/2, u + 1 + k/2)$ closes a two-star to form a triangle. Hence, finding the number of triangles per node reduces to evaluating the grey area marked in this figure, which is found to be $\triangle = (3/2) \times ((k/2)^2 - k/2)$.

A special case of the k-regular nearest neighbor lattice is the complete network ($k = N - 1$). In this network of N nodes, every possible link exists. It is therefore a model of perfect interconnectedness with an average shortest path length of $\langle s \rangle = 1$ and unit clustering $C = 1$. Due to the fact that every node is connected to every other node, the complete network is an ideal model for perfectly well-mixed systems as a substrate for dynamic processes: At any time, each node has complete information about the state of the whole system. In contrast, a node in a k-regular nearest-neighbor lattice has only access to its local environment. An example complete network is shown in Fig. 2.14.

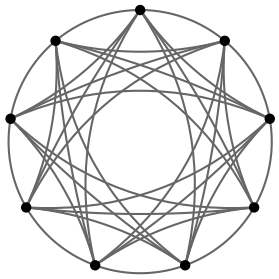


Figure 2.14: A special case of a k-regular network is the complete network in which each node is connected to each other node (and therefore has degree $k = N - 1$). Depicted here is an example with $N = 9$.

In order to investigate mixed systems with higher sparsity and less clustering, the random k-regular network model is used. Within this model, each node has exactly k neighbors, but connections between nodes are chosen at random. Since the connections are random, the clustering coefficient is approaching $C \to 0$ for $k \ll N$. An efficient sampling algorithm is given in [72].

Within well-mixed systems of lower density, it is assumed that any two entities have uniform probability of encountering each other within a fixed amount of time and that any entity has access to a similar-sized subvolume of the system. In physical chemistry, this approximation is often used to model the temporal evolution of the number of chemical reactants given their chemical reaction equa-

tions [73]. The k-regular network, due to its definition, represents a "quenched disorder" state of such a system where connections between entities are sampled once and then remain. It is therefore a useful structure to study how the discretization of a contact system changes the outcome of certain dynamic processes taking their course on the substrate compared to the well-mixed case. For an explicit example of this, see the discussion of epidemic spreading in Sec. 3.2. It will further be used as a model network for which a new heuristic to compute the cover time of random walks yields almost exact results (see Ch. 6).

2.2.2 *Erdős–Rényi Random Networks*

One of the simplest stochastic network models is defined as follows: In a network of N nodes, connect each pair of nodes (u, v) with probability p_{ER}.[8] The ensemble defined by the number of nodes N and this connection probability p_{ER} is called the $\mathcal{G}(N, p_{\mathrm{ER}})$ *random graph.* Even though this definition was introduced by Gilbert in 1959 [75], it is widely attributed to the names of Erdős and Rényi, who defined a similar ensemble $\mathcal{G}(N, m)$ in the same year [76]. In their model, a fixed number m of all node pairs is chosen to be connected. However, in this thesis, only the $\mathcal{G}(N, p_{\mathrm{ER}})$ model will be used (due to their similarity it suffices to discuss one of them) but referred to as the Erdős–Rényi (ER) random graph.

[8] While this description suffices to define a method to sample networks from the ensemble, a more efficient sampling algorithm is given in [74]. If not stated otherwise, the latter algorithm will be used in this work whenever Erdős–Rényi networks are constructed.

The random graph is widely used as a reference model to find structural features of other network models and empirically observed static networks which are not caused by noise. For instance, the clustering coefficient is relatively large in social networks (e.g. $C = 0.51$ in the author's Facebook friend network), whereas the mean degree is small compared to the number of nodes (an upper bound is given by e.g. Dunbar's number which places the maximum possible number of active social relationships per individual at ≈ 150 [77]). Following the definition in Sec. 2.1.4, the clustering coefficient is the probability that two neighbors of a node are connected. Since in the Erdős–Rényi model every edge is independently drawn with probability p_{ER}, the clustering coefficient is simply equal to this probability

$$C_{\mathrm{ER}} = p_{\mathrm{ER}}. \tag{2.7}$$

Furthermore, a single node has $N - 1$ potential neighbors, each of which it connects to with probability p_{ER} and so its mean degree will be

$$\langle k \rangle = p_{\mathrm{ER}} \times (N - 1). \tag{2.8}$$

Assuming that the mean degree remains constant with growing number of nodes N, the connection probability consequently vanishes and so does the clustering coefficient. Hence, the ER random graph model is usually not a proper candidate to model social networks. It serves as a reference model to networks of similar mean degree.

One can further compute the complete degree distribution. Because a single node has $N - 1$ potential neighbors and connects to

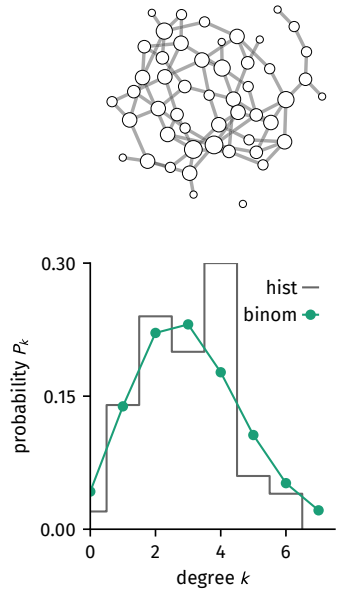


Figure 2.15: A sample Erdős–Rényi network and its degree distribution for parameters $N = 50$ and $\langle k \rangle = 3$, compared to the theoretical binomial distribution Eq. (2.9).

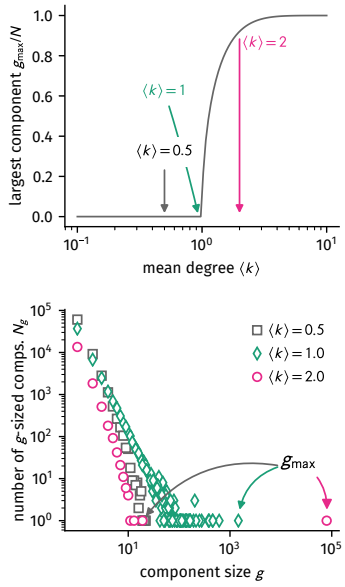


Figure 2.16: Largest component and component size distribution on samples of the ER network with $N = 10^5$. Below the critical value of $\langle k \rangle_c = 1$ components form whose sizes are distributed exponentially. The relative size of the largest component $g_{\max}/N$ vanishes, however. At the critical point, the component size distribution follows a power-law. Above the critical value, a large component of size $\mathcal{O}(N)$ forms. The size distribution of the remaining components follows an exponential shape again.

each of those with probability p_{ER}, the probability that a node has degree k is equal to the probability of k "successes" and $N-1-k$ "failures" in $N-1$ Bernoulli trials which is

$$P_k = \binom{N-1}{k} p_{\text{ER}}^k (1-p_{\text{ER}})^{N-1-k}. \tag{2.9}$$

An example is shown in Fig. 2.15. Node degrees of the ER graph therefore follow a binomial distribution $k \sim \mathcal{B}(N, p_{ER})$ with variance

$$\text{Var}[k] = N p_{\text{ER}}(1-p_{\text{ER}}) \overset{N \gg 1}{=} \langle k \rangle \ , \tag{2.10}$$

which implies that nodes are rather homogeneous concerning their number of connections. An example is shown in Fig.Another important property of ER networks is that nodes are quickly reachable from all over the network. The average shortest path length scales as

$$\langle s \rangle \propto \log N,$$

which is substantially lower than in, e.g. k-regular nearest neighbor lattices. The largest component of the Facebook friend network has a shortest path length of $\langle s \rangle = 3.6$ and is composed of $N = 329$ nodes which yields $\log N = 2.5$ for the ER network model, which is of similar order. The ER model therefore accurately reproduces short paths in the network, however, its vanishing clustering coefficient fails to accurately model the Facebook network's clustering of $C = 0.51$.

Like random k-regular networks, Erdős–Rényi networks are particularly useful as a first step towards discrete systems when analyzing "well-mixed" systems, especially because they are easily tractable analytically due to their simple definition. Bond percolation is an example [32, pp. 347 ff.]. In percolation problems, one aims at finding the critical density of either sites (in networks: nodes) or bonds (in networks: links) at which a giant component emerges: A component which consists of a non-vanishing number of nodes in the thermodynamic limit $N \to \infty$. Before this critical density, the network consists of components whose sizes are distributed exponentially. The largest component of size $g_{\max}$ will vanish relatively to the number of nodes $g_{\max}/N \to 0$. At the critical point, a phase transition occurs; the component size distribution will follow a power-law. For densities above the critical value, a giant component of size $g_{\max}/N > 0$ exists. The size of the remaining components is distributed according to an exponential distribution again. For the ER network model and link percolation, this critical density is given as $p_{\text{ER},c} = 1/(N-1)$, i.e. at $\langle k \rangle_c = 1$ (see [32, pp. 347 ff.]). One may find this result intuitive: each average node has to connect to more than one other node in order for structures larger than pairs to form. The group-size distribution and average largest component size of the ER network model are illustrated in Fig. 2.16.

In this work, the ER model will be used to study how the first passage time of random walks changes when varying the substrate structure between lattices and the random graph in multiple ways

(Part II). A temporal version of the ER graph (introduced in Sec. 2.4.2) will be used to study the spread of diseases in temporal networks in Part III. In this context, the critical degree for a large component to form will be of importance.

2.2.3 *Random Geometric Graphs*

A two-dimensional random geometric graph (RGG) [78] can be interpreted as a random model for simple two-dimensional lattices, where nodes are connected to nodes in their vicinity. By varying the definition of 'vicinity' (i.e. the connection radius), the model can be used to model the spread of infectious diseases with varying reach in groups of humans [79].

RGGs can be defined in the following way. Distribute N nodes on the unit square $[0,1)^2$ where their positions are drawn randomly from a uniform distribution as $x_u, y_u \sim \mathcal{U}(0,1)$ for each node u. Here, $\mathcal{U}(a,b)$ refers to a continuous distribution of values on the inverval $[a,b)$ with constant probability density $f(u) = 1/(b-a)$. Let $\boldsymbol{r}_u = (x_u, y_u)$ be the position vector of any node u. Then, an edge between a pair of nodes (u,v) is drawn if and only if $r_{uv} = |\boldsymbol{r}_u - \boldsymbol{r}_v| \leq R$ where R is the connection radius. These networks have a homogeneous degree distribution due to the fact that the probability of a node having k neighbors is approximately equal to the probability of exactly k nodes being inside a disc with radius R around the focal node which is of Poissonian shape. As derived in [78], in the limit of large N, the clustering coefficient is given as

$$C_{\text{2DRGG}} = 1 - \frac{1}{\sqrt{\pi}} \times \frac{1}{\Gamma(3/2)} \times \left(\frac{3}{4}\right)^{3/2}, \tag{2.11}$$

where "2DRGG" implies "two-dimensional random geometric graph". In the equation above, $\Gamma(x)$ denotes the Gamma function. Note that this result does not depend on the connection radius R. As RGGs are modeling lattices, their average shortest path length is typically higher compared to the average shortest path length of, e.g. Erdős–Rényi networks. An example RGG is shown in Fig. 2.17.

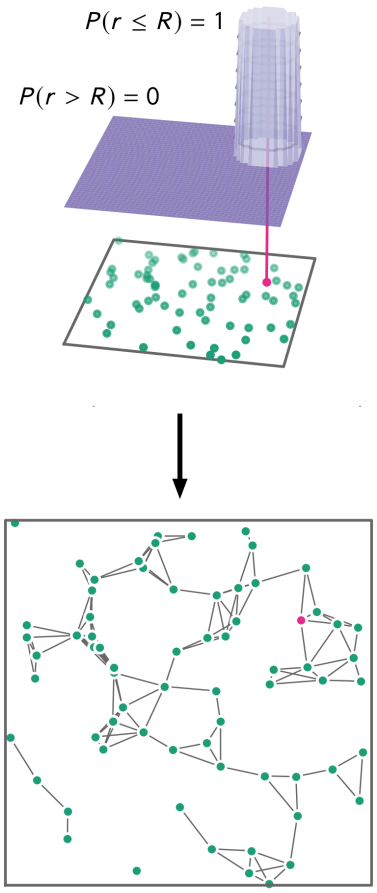


Figure 2.17: A sample random geometric graph in two dimensions without periodic boundary conditions.

2.2.4 *Small-World Effect and Corresponding Network Models*

As mentioned several times before, social networks carry two very important properties. First, they are highly clustered. An illustration was given by the author's Facebook friend network, in which the global probability that two neighbors of a node are neighbors, too, is $C = 0.51$. The second fundamental property is that they are 'small-world', which means that their average shortest path length scales as $\langle s \rangle \propto \log N$ [1]. The emergence of short paths between two unconnected nodes in social systems was first demonstrated in Travers' and Milgrim's small-world experiment [8]. Within their study, the researchers handed 300 packets to individuals spread across the U.S. and gave them the task to send the packet to a single person located in Boston. This assignment came with a caveat: An individual was only

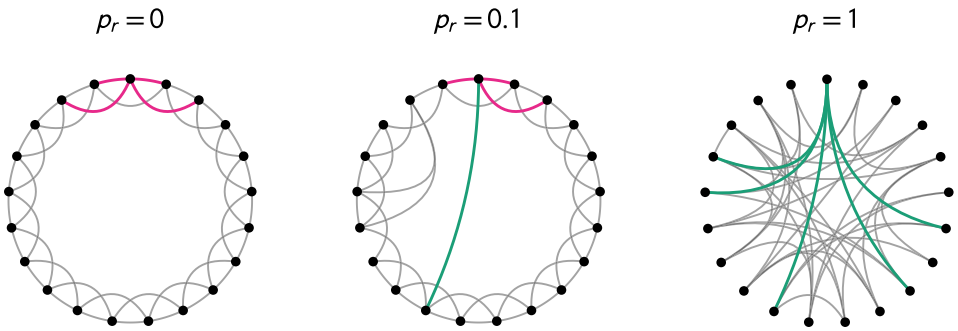


Figure 2.18: Some example networks constructed with the algorithm by Watts and Strogatz [1]. At rewiring probability $p_r = 0$, the networks are equal to a k-regular nearest neighbor lattice. With $p_r = 0.1$, roughly 10% of the edges have been rewired, possibly connecting a few nodes to other nodes further away, creating short paths. At $p_r = 1$, every edge has been rewired, such that the resulting network is strongly randomized (but not equal to an ER network).

allowed to send the packet to the target person if they knew them on a first-name basis. If they did not, they had to send the packet to another person they did know on a first-name basis, hoping this person would know the target person. This second person then had to act in the same way. Roughly 20% of the packets arrived at the target person, each of those trough chains of an average length of ≈ 6. It would be foolish to assume that this number is equal to the average shortest path lengths of the acquaintance network in the U.S.—after all, ca. 80% of the packets were lost, the source sample was biased (≈ 100 source persons lived near Boston and ≈ 200 source persons lived in Nebraska), and the reported chains were by no means assured to be the shortest. Nevertheless, it was surprising that such short chains exist at all. In order to model social structures which are clustered and spread out, but still possess short chains, a network modeling approach by Watts and Strogatz connected two of the simple network models introduced above [1]. ER networks were shown to possess low average shortest path lengths, but do not possess any clustering. On the other hand k-regular networks show high clustering but walkers require, on average, substantially more steps to travel from one node to another. Watts and Strogatz construct small-world social networks based on those models as follows.

One begins with a k-regular nearest neighbor lattice on a ring and introduces a structural control parameter, the rewiring probability p_r. For each node, one iterates through each of its $k/2$ rightmost neighbors. With probability p_r, each edge connecting the focal node to the corresponding rightmost neighbor is cut. Afterwards, a second node is chosen with the condition that this second node is different from the focal node, the former rightmost neighbor, and any other node the focal node is already connected to. Then, the edge is rewired between the focal node and the chosen new neighbor.

Examples of resulting structures for different values of p_r are shown Fig. 2.18. In the limit $p_r = 0$, no edge rewires and therefore the network is equal to the k-regular nearest neighbor network. Clustering is high and the average shortest-path length is, as well. For growing rewiring probability, a small amount of edges is rewired to

nodes further away and hence, the average shortest path length decreases, while only a small number of triangles is broken. At $p_r = 1$ every edge has been randomized. Therefore clustering is low and so is the average shortest path length. Systematically increasing the rewiring probability p_r yields a picture illustrated in Fig. 2.19a. Already with a small amount of rewired edges, the average shortest path length decreases rapidly, while global clustering persists. At the same time, the mean degree stays constant, but the degree distribution becomes more heterogeneous (see Fig. 2.19b).

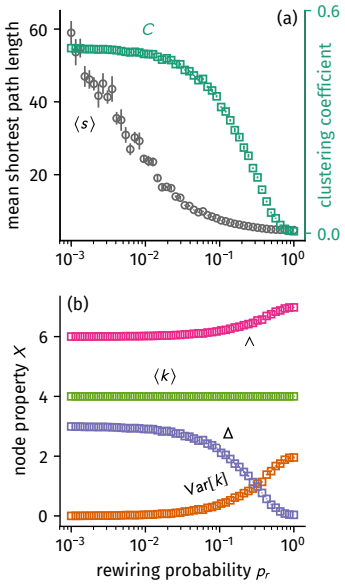


Figure 2.19: Properties of Watts–Strogatz networks. With only a few rewired edges (low p_r), the average shortest path length $\langle s \rangle$ decreases rapidly, while the global clustering C persists. The mean degree stays constant because the number of edges remains constant. However, the degree distribution becomes broader (increasing Var $[k]$).

This result conceptualizes the following: In a world where the population is spread out and only knows its local surroundings, it suffices to have a few long-range connections for short communication paths to appear.

A second, more complicated problem is associated with the question how people are able to identify these short chains to send messages efficiently. Certainly, they do not have knowledge about the whole network. To answer this question, Kleinberg introduced a variation of a small-world model with an inherent notion of distance [31], see Fig. 2.20. Within this model, nodes are positioned according to a two-dimensional lattice and connected to their nearest neighbors. Furthermore, the notion of distance d_{uv} between two nodes u and v is introduced by setting it equal to the shortest path length s_{uv} on the lattice. Afterwards, two nodes u and v are connected with a probability algebraically decaying with distance as $p_{uv} \propto d_{uv}^{\mu}$. The exponent $\mu \leq 0$ therefore controls the amount of long-ranged edges added to the underlying structure of local connections. Kleinberg revealed that under certain conditions there exists an optimal value $\mu = -2$ at which a lower bound of the message delivery time (measured in steps) becomes minimal. The conditions are based on Travers' and Milgrim's experiment and read as follows:

Any node u currently carrying a message which is supposed to be delivered to a target node t has information about

> "(i) the set of local contacts among all nodes (i.e. the underlying grid structure);
>
> (ii) the location, on the lattice, of the target t; and
>
> (iii) the locations and long-range contacts of all nodes that have come in contact with the message" [31, p. 165].

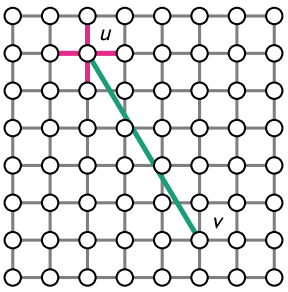


Figure 2.20: Illustration of the Kleinberg small-world model [31]. The basis is a two-dimensional lattice where each node u is connected to its nearest neighbors, where the distance d_{uv} between two nodes u and v is given by their shortest path length on the lattice. The focal node is furthermore connected to every other node v with some distance-dependent probability $p_{uv} \propto d_{uv}^{\mu}$, $\mu \leq 0$.

These conditions imply that individuals are informed about the spreading process itself, the local structure of the network and some have a notion of distance. Yet, natural spreading processes like diffusion, synchronization of oscillators, or spreading of diseases may follow more random, non-informed behaviors [34]. It will therefore be of interest to compute message passing times of random processes in such small-world networks.

One particular network property which is not reflected by any of the two models introduced above is that networks are often organized in a modular hierarchical structure (see Sec. 2.1.7). This problem was approached with the introduction of a third small-world

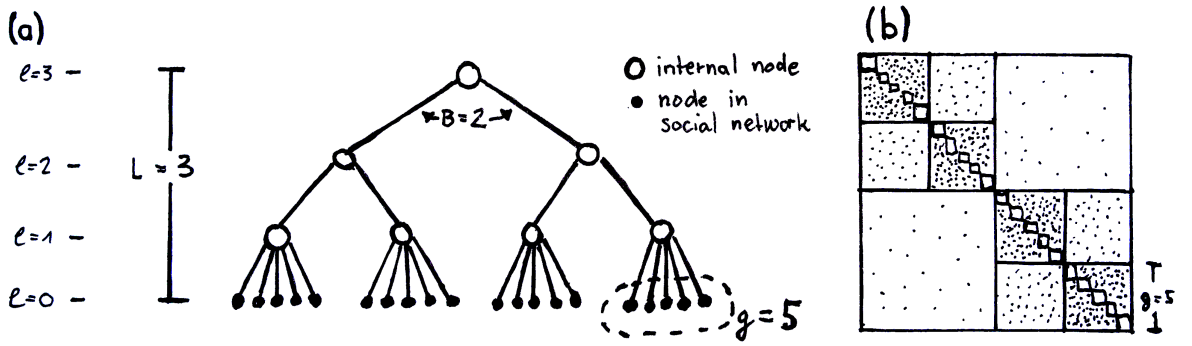


Figure 2.21: Illustration of the modular hierarchical social network modeling approach presented in [13]. **(a)** Modular hierarchical community structure is introduced by considering a modified B-ary tree of height L. Each internal node (represented by ○) groups together either B subgroups or g leaf nodes. Hence, each leaf node of the final social network (represented by •) is locally part of a group of size g. The social distance between two nodes is given by the distance of their lowest common ancestor in this hierarchical tree. **(b)** A schematic average adjacency matrix as produced by the tree shown in (a). Darker shades represent a higher connection probability. Due to this structure, such models are often referred to as stochastic block models [80, 81].

network model by Watts, Dodds, and Newman [13]. Within this model, each node is part of a local group of size g which is organized in a hierarchical tree (see Fig. 2.21a). Each pair of nodes (u, v) is associated with a hierarchical distance ℓ_{uv}, which is equal to the number of steps to find their lowest common ancestor in the tree. Nodes in the same group have hierarchical distance $\ell = 1$, nodes whose groups are distinct but lie in the same module are of distance $\ell = 2$ and so forth. Then, two nodes are connected with probability $p_\ell = c \exp(-\alpha(\ell - 1))$ which produces a nested-block structure in the average adjacency matrix (see Fig. 2.21b). Here, the control parameter α controls the amount of 'long-distance' connections a node has where 'distance' refers to a social distance. It is not a distance in the classical sense because the triangle inequality does not hold. One particular advantage of this model is that at $\alpha = 0$, the model is equal to the ER network model, which is not the case for the other network models. Furthermore, the mean degree is kept constant which means that on average, the number of connections in the system remains constant, regardless of the chosen value of α—this is the case for the Watts–Strogatz model but not for the Kleinberg model. Using this hierarchical model, the authors showed that by the same informational conditions that Kleinberg gave, there exists a value of the control parameter α at which the message passing time becomes minimal. They further relate that the Kleinberg model is similar to theirs. For their proof, they do not use the explicit network structure of this model but only work with connection probabilities between nodes of distance ℓ_{uv}. Hence, their results are valid for the average medium but do not necessarily reflect the properties of actual single network samples.

Based on the models introduced in this subsection, this thesis is going to investigate the following points. First, these models are all related in their concepts, however, two of them do not reach the ER model in the randomized limit, which might lead to difficulties when comparing dynamic systems. Furthermore, it is unclear how exactly they can be mapped to each other, such that their network properties can be compared. In Ch. 4, alternative models are introduced which

all reflect the original models but come with the advantage that their structural control parameters can be directly mapped to each other and their randomized limit will represent the ER model. This allows for a direct comparison of their network properties and the definition of a generalized small-world model. It will further be shown that the alternative models greatly ease the analytical evaluation of certain observables. Using these models, it will be investigated in which way true modular hierarchical structure differs from the other small-world structure and how it influences dynamics, i.e. the spread of random walkers (Ch. 5 and Ch. 6). Using a random message passing approach, it will be shown that certain small-world models exhibit optimal structure to efficiently forward messages even without knowledge about structure and the spreading processes. This will be revealed to be a pure network effect which cannot be explained without taking into account the explicit connectivities of nodes. It will further be shown that this effect vanishes in another model where the nodes' connectivities remain constant on average.

2.2.5 *Barabási-Albert Networks*

Real-world networks usually possess another important property which has been left unmentioned so far, albeit hinted at in the analysis of the author's Facebook friend network: Node degrees are often distributed according to heavy-tailed distributions, which means that most nodes have a low number of connections and a very small number of nodes have very high numbers of connections. These so-called hubs are responsible for a very short average path length and influence a number of dynamic processes drastically.

In order to model degree heterogeneity, Barabási and Albert devised a random network model that produces networks with heavy-tailed degree distributions [82]. Within the model, a small seed network is considered initially. Afterwards, nodes enter the network incrementally until the desired number of nodes N is reached. Each time a node enters, it picks m other nodes v to which it connects, each of which is chosen with probability $p_v \propto k_v$. In [83] it has been shown that this produces networks of degree distribution

$$P_k = \frac{2m(m+1)}{k(k+1)(k+2)}, \tag{2.12}$$

which scales as $P_k \propto k^{-3}$ for large degrees. Otherwise, the networks possess random structure such that, for instance, the clustering coefficient vanishes.

Barabási-Albert networks will be used to illustrate epidemic spreading in systems of heterogeneous degree distribution in Sec. 3.2.3 and as a model on which a newly found heuristic is able to compute the random walk cover time with high accuracy (see Ch. 6).

Strong heterogeneity in the degree distribution greatly influences the behavior of spreading processes as will be shown in Ch. 3. Therefore, any modeling of small-world and modular hierarchical networks will omit such heavy-tailed degree distributions in order to

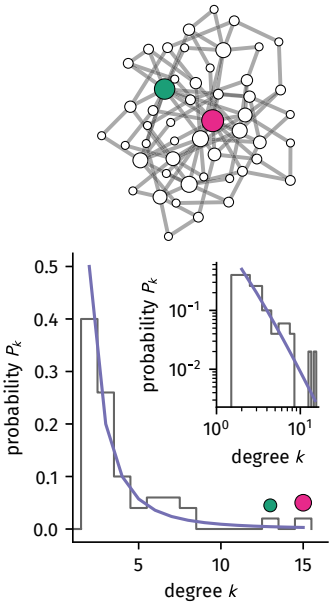


Figure 2.22: A sample Barabási-Albert network of $N = 50$, $m = 2$ and its degree distribution. Two hubs are observed (nodes with large degree, colored teal and pink). The degree distribution roughly follows the theoretical prediction Eq. (2.12).

avoid the obfuscation of the small-world structure's influence on spreading dynamics.

2.3 Temporal Networks

Networks do not necessarily have to be static. In fact, many static networks are derived from data collected over a certain period of time, i.e. they represent time-aggregated temporal data [35, 84]. Examples include social networks of humans or animals where contacts between individuals are recorded for a certain period of time (e.g. [12, 39, 40, 85]), or transportation networks, where in a static picture, edges represent possible connections between stations or stops but actual traversing between those may only occur on times when a train or a bus runs (see e.g. ref. [84], Sec. 2.8). In recent years, interest in analyzing the original temporally resolved data has increased as opposed to analyses of the aggregated static network. The former has been done in the hope of simplifying analyses by the increased amount of information considering the temporal dimension. Naturally, since temporal aggregation projects data on a lower-dimensional space, information about how a node pair was in contact explicitly is lost when producing aggregated static networks. It has since been found, that, e.g. the inference of community structure is simplified on temporal network data [86], or that during the spread of a disease, not every node is accessible for an infection when respecting the arrow of time in contact data [87], as opposed to a spreading disease on the corresponding aggregated network. On the other hand, a temporally aggregated network of air traffic passengers between airports suffices to reliably predict the origin and arrival time of a globally spreading disease [3, 4]. In the context of temporal human proximity networks, a central question is how the 'temporality' of a network influences the spread of a disease, where 'temporality' is ill-defined, but the question might be interpreted to roughly mean "how different is the representation of the system as a temporal network compared to an aggregated network for a spreading disease". The aim of this thesis will be to introduce and analyze a new model that replicates certain properties of empirical temporal human proximity networks and to investigate the implications of this model's properties on the outcome of epidemic spreading.

In this section, temporally resolved human proximity networks and their properties are formally defined and three data sets of empirically recorded human proximity networks are introduced. A collection of important structural and temporal observables of these data sets are measured and described in detail.

2.3.1 Temporal Contact Networks

The temporal networks discussed in this work will describe temporally varying face-to-face contacts between a finite set of human individuals for the purpose of investigating epidemic spreading. It

suffices to consider unweighted, undirected networks with constant node number N. An edge between a node u and a node v means that u can infect v and vice versa. Hence, the temporally varying structure of a network will be quantified by a temporally varying $(N \times N)$-sized adjacency matrix

$$A_{uv}(t) = \begin{cases} 1 & \text{if } u \text{ and } v \text{ are connected at time } t, \\ 0 & \text{otherwise.} \end{cases}$$

In this adjacency matrix, contacts change as step functions in continuous time. Two ways to describe the changing structure of the temporal network will be introduced below, illustrated in Fig. 2.23. [9]

[9] In the following, nodes are considered to be denoted as integers with the condition $0 \leq u < N$ for each node u.

A node pair (u, v) is associated with an ordered set of time pairs $(t_{i,a}^{(u,v)}, t_{i,d}^{(u,v)})$ where $t_{i,a} < t_{i,d} < t_{i+1,a}$. The subscript "$a$" denotes an activation time and the subscript "d" denotes a deactivation time. For each time pair i, this means that the edge is 'turned on' (exists and connects u and v) during the half-open time interval $[t_{i,a}^{(u,v)}, t_{i,d}^{(u,v)})$. It is 'switched off' (does not connect u and v) in the time intervals $[t_{i,d}^{(u,v)}, t_{i+1,a}^{(u,v)})$. Furthermore, temporal networks will be considered to have been collected during an experiment: At some time t_0, the experiment began and edges were recorded. Therefore, edges which exist at time t_0 may have existed before, such that the contact started at an unknown, unrecorded time. Nevertheless, if this is the case for a pair (u, v), the first activation time will be set to $t_{0,a}^{(u,v)} = t_0$. Note that $t_0 \leq t_{i,a}^{(u,v)}$ for all node pairs (all recorded activation events are of later time or equal to the time at which the recording process began). Furthermore, the data collection process is assumed to have ended at a time $t_{\max}$ such that the whole observation period has length $T = t_{\max} - t_0$. Edges that have not been turned off before $t_{\max}$ may exist longer—the exact deactivation time is unknown. Its deactivation time will nevertheless be set to $t_{i,d}^{(u,v)} = t_{\max}$. Therefore, $t_{i,d}^{(u,v)} \leq t_{\max}$ for all node pairs.

One may group certain activation and deactivation times together if they are equal for multiple edges, resulting in a chain of global event times $t_0 < t_1 < t_2 < \cdots < t_{\max}$. The adjacency matrix changes at the exact event times t_j. At time t_0, the network consists of an initial edge set of activated contacts

$$\varepsilon_0 = \left\{ (u, v) \middle| 0 \leq u < v < N \text{ and } t_{0,a}^{(u,v)} = t_0 \right\}.$$

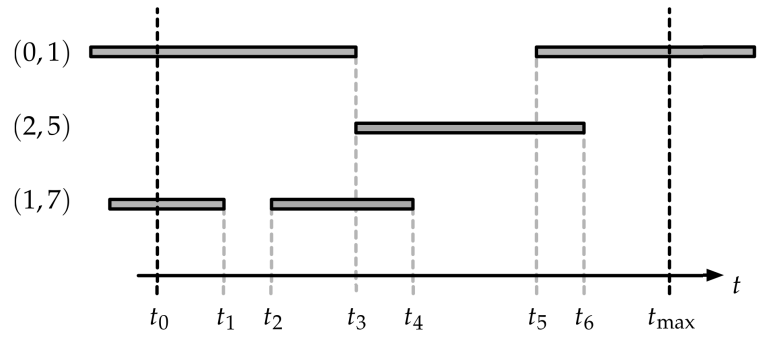
At each event time $t_0 < t_j < t_{\max}$, the network change is quantified by two edge sets. The first one contains the edges 'leaving' the network (being deactivated) such that

$$\varepsilon_j^{(\text{out})} = \left\{ (u, v) \middle| 0 \leq u < v < N \text{ and } t_j \in \{ t_{0,d}^{(u,v)}, t_{1,d}^{(u,v)}, \ldots \} \right\}.$$

The second set contains all edges 'entering' the network (being activated) with

$$\varepsilon_j^{(\text{in})} = \left\{ (u, v) \middle| 0 \leq u < v < N \text{ and } t_j \in \{ t_{0,a}^{(u,v)}, t_{1,a}^{(u,v)}, \ldots \} \right\}.$$

(a)



(b)

node pair	activation time pairs
$(1,7)$	$\{(t_0,t_1),(t_2,t_4)\}$
$(2,5)$	$\{(t_3,t_6)\}$
$(0,1)$	$\{(t_0,t_3),(t_5,t_{\max})\}$

(c)

time	initial edge set ε_0
t_0	$\{(0,1),(1,7)\}$

time	edges leaving $\varepsilon_j^{(\text{out})}$	edges entering $\varepsilon_j^{(\text{in})}$
t_1	$\{(1,7)\}$	$\emptyset$
t_2	$\emptyset$	$\{(1,7)\}$
t_3	$\{(0,1)\}$	$\{(2,5)\}$
t_4	$\{(1,7)\}$	$\emptyset$
t_5	$\emptyset$	$\{(0,1)\}$
t_6	$\{(2,5)\}$	$\emptyset$

Figure 2.23: Illustration of an example temporal network with $N = 8$ nodes recorded between times t_0 and $t_{\max}$ and two ways how the changing structure can be recorded. **(a)** illustration of how links change over time in a temporal network. Each node pair (u,v) is depicted as a row. A grey horizontal bar means that the edge between u,v was active during this time (i.e. nodes u and v are connected). Shown here are only node pairs which were active during the observation period. Note that the links $(0,1)$ and $(1,7)$ were active before the observation period began and the observation was halted while link $(0,1)$ was still active. Therefore, the true first activation times of $(0,1)$ and $(1,7)$ are unknown, as well as the true last deactivation time of $(0,1)$. **(b)** the changing network structure as a collection of time pairs in which each link was active. **(c)** the changing network as an initial edge list and a chronologically ordered list of edge changes.

Note that due to the definitions above, these edge sets are necessarily distinct such that

$$\varepsilon_j^{(\text{out})} \cap \varepsilon_j^{(\text{in})} = \emptyset.$$

At the end of the experiment at time $t_{\max}$, no edge changes are recorded.

Due to these definitions, at any event time $t_0 < t_j$, the edge set is updated as

$$\varepsilon(t_j) = \varepsilon(t_{j-1}) - \varepsilon_j^{(\text{out})} + \varepsilon_j^{(\text{in})},$$

with $\varepsilon(t_0) = \varepsilon_0$. This edge set completely describes the temporal network during the half-open time interval $[t_j, t_{j+1})$ where $t_{j+1} \leq t_{\max}$.

It is common practice to discretize time when working with temporal networks, which is either done by 'binning' or 'sampling'. In this work, temporal networks will be treated as objects of continuous-time. Nevertheless, the temporal network data used in this thesis was recorded in a binned manner, and some studies refer to sampled networks. Therefore, both will be introduced here.

When changing from a continuous-time to a discrete-time perspective, the observation period $[t_0, t_0 + T)$ is divided into $T/\Delta t$ equal-sized time intervals of length Δt, such that the new event times are given as $\tilde{t}_j = n\Delta t$ with $n \in \mathbb{N}^+$, $0 < n < T/\Delta t$.

In a *binning* procedure, if an edge existed at any time within the time interval $[n\Delta t, (n+1)\Delta t)$, it is set to have existed during the whole time interval. Effectively, this extends a contact's duration by decreasing its activation time and increasing its deactivation time

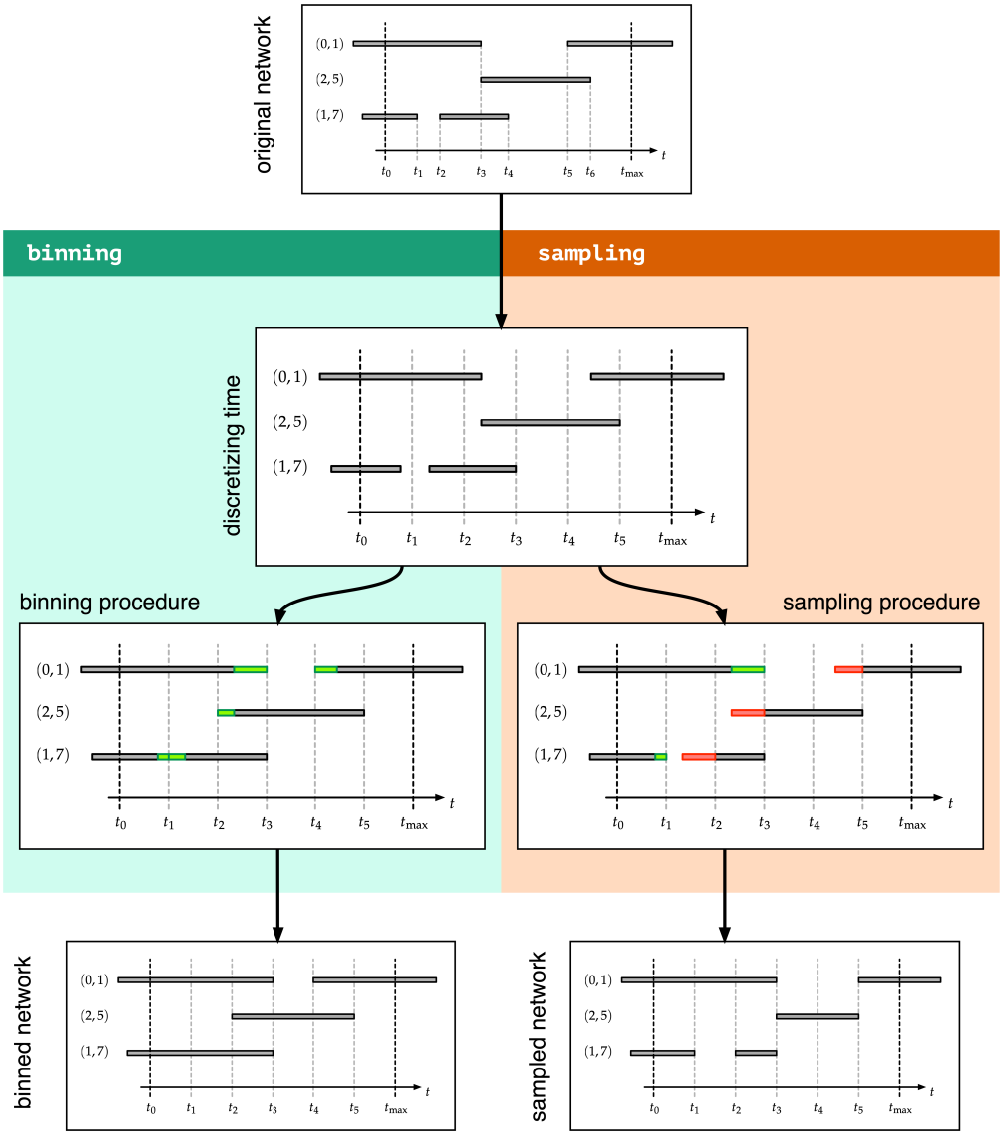


Figure 2.24: Discretizing time by binning and sampling a temporal network. During a binning procedure, an edge which has been active during a single discrete time interval is interpreted to have been active for the whole time interval. Formerly not overlapping edges may overlap. The active duration of edges which were shortly inactive may be artifically prolonged. During a sampling procedure, edges which are active during the beginning of a discrete time interval are assumed to be active for the whole time interval. Edges which were not active at the beginning of the interval are assumed to not have been active for the whole time interval. Short contacts may be erroneously overlooked by this procedure.

(see Fig. 2.24), potentially producing paths which did not exist in the system by creating overlaps of edges which share nodes.

When *sampling* a temporal network, the original edge set is sampled every Δt. This edge set is consequently assumed to be active for the whole time interval $[n\Delta t, (n+1)\Delta t)$ (see Fig. 2.24). Doing so potentially increases a contact's activation time to the beginning of the following interval. It also increases a contact's deactivation time to the end of the corresponding sampling interval. Contacts of duration smaller than Δt may potentially be omitted, missing possible transmission pathways.

Both procedures yield the original network only for $\Delta t \to 0$. Note that single network samples that each exist during a single discrete time-step will be referred to as *snapshots*.

2.3.2 Temporal Network Observables

Basic Properties

Some simple structural and temporal properties of temporal networks quantify the system's density and activity over time.

A node's degree at time t is given by Eq. (2.2) where the adjacency matrix is replaced by the time-dependent adjacency matrix. The mean degree

$$\begin{aligned}\langle k(t)\rangle &= \frac{2}{N(N-1)}\sum_{u=1}^{N-1}\sum_{v=u+1}^{N} A_{uv}(t)\\ &= \frac{2}{N}\left|\varepsilon(t)\right|\end{aligned}$$

is proportional to the number of contacts in the system at time t and is therefore an observable which quantifies the current structural state of the network.

The number of edges being activated and deactivated during a time-interval $(t, t+\Delta t]$ are given as

$$e_{\text{in}}(t, t+\Delta t) = \sum_{j=1}^{j'} \left|\varepsilon_j^{(\text{in})}(t)\right| \tag{2.13a}$$

$$e_{\text{out}}(t, t+\Delta t) = \sum_{j=1}^{j'} \left|\varepsilon_j^{(\text{out})}(t)\right| \tag{2.13b}$$

where j refers to the smallest event time where $t < t_j$ and j' refers to the largest event time at which $t_{j'} \leq t + \Delta t$. These observables quantify the activity of the network during a time interval—the amount of structural change a network undergoes in a certain period of time. Note that in contrast to the definitions above, the half-open interval includes the upper bound and excludes the lower bound here, which is done in order to associate the first change in the network with the first interval for interval sizes $\Delta t \geq t_1 - t_0$.

Over the whole observation period, the network may be in several different structural states. In order to quantify a whole-observation picture, the time-averaged degree distribution will be discussed, which is given as

$$\overline{P_k} = \frac{1}{T}\int_{t_0}^{t_{\max}} dt\ P_k(t).$$

In the context of epidemic spreading, nodes may be infected for a long time. In this case, they 'see' all of the network before recovering. In this limit, the averaged network is allegedly sufficient to describe the interactions between nodes appropriately. The average network is given as

$$\overline{A_{uv}} = \frac{1}{T}\int_{t_0}^{t_{\max}} dt\ A_{uv}(t).$$

Note that sometimes this will be referred to as the aggregated network, because the difference between the averaged network and the

aggregated network is merely a factor T in every entry of the weighted adjacency matrix, such that there is no structural difference between the two.

Inter-Contact and Contact Durations

In order to quantify the behavior of human beings, the following simple temporal observables are defined. In the context of epidemic spreading, a contact between a susceptible and an infected node may yield an infection event. The probability with which such an infection happens increases with increasing time. Therefore, one observable of interest is a contact's duration τ_C. Each activation time pair i of a node pair (u, v) is associated with a contact duration

$$\tau_C = t_{i,d}^{(u,v)} - t_{i,a}^{(u,v)} \qquad \forall t_{i,a} \neq t_0 \text{ and } t_{i,d} \neq t_{\max}.$$

Note that following this definition, contacts which existed at the beginning of the observation or lasted until the end of the observation are omitted because their true duration is unknown.

Second, isolated nodes may recover without infecting any other node, the probability of which increases with increasing inter-contact duration—the total amount of time a node spends isolated between their last and their next contact. Phenomenologically, the time a node spends alone is equal to the duration of a group of size $g = 1$. Therefore the definition of the inter-contact duration is equivalent to the following definition of group life-time for groups of any size.

Group Sizes

The empirical network data introduced in the following section will show to have an outstanding property: At any point in time, face-to-face contact networks consist of a collection of densely connected, distinct components, referred to as *groups* in the following. One major contribution of this work will be the discussion of this group-structure's influence on epidemic spreading. To this end, several group-related observables will be introduced here.

In analogy to Sec. 2.1.5, the group-size distribution of a network at time t is quantified by the number $n_g(t)$ of g-sized components. The procedure described in the following sequentially updates the component sets and measures the duration τ_g of a group of size g.

Let $\mathcal{G}_j$ be the set of all components of the initial network defined by the edge set ε_j for each time $t_0 \leq t_j$, each evaluated by running the algorithm described in Sec. 2.1.5. Each component $\kappa_\ell^{(j)} \in \mathcal{G}_j$ is a set of nodes with the properties

$$\kappa_\ell^{(j)} \cap \kappa_{\ell'}^{(j)} = \begin{cases} \emptyset & \text{if } \ell \neq \ell' \\ \kappa_\ell^{(j)} & \text{otherwise, and} \end{cases}$$

$$\bigcup_{\ell=1}^{|\mathcal{G}_j|} \kappa_\ell^{(j)} = \{u | u \in \mathbb{N}, 0 \leq u < N\}.$$

At time t_0, each node u of the network is associated with the activation time $t_a(u) = t_0$ and the component set $\mathcal{G}_0$ with $\kappa^{(0)}(u)$ denoting the component in which node u resides initially. Then, chronologously iterating through each edge change event at time $t_0 < t_j < t_{\max}$, update the network according to $\varepsilon_j^{(\text{in})}$ and $\varepsilon_j^{(\text{out})}$, evaluate the new set of components $\mathcal{G}_j$ and denote as $\kappa^{(j)}(u)$ the component in which u resides after the update. For each event j, iterate through every component $\kappa^{(j)} \in \mathcal{G}_j$. For a single node u in $\kappa^{(j)}$, check if $\kappa^{(j)}(u) = \kappa^{(j-1)}(u)$. If this is the case, the group persisted trough the update, hence nothing has to be done. If $\kappa^{(j)}(u) \neq \kappa^{(j-1)}(u)$, the group $\kappa^{(j-1)}(u)$ has changed. In this case, the duration of the g-sized group with $g = \left|\kappa^{(j-1)}(u)\right|$ is

$$\tau_g = t_j - t_a(u).$$

Afterwards, set the activation time of each node $v \in \kappa^{(j)}(u)$ to $t_a(v) = t_j$ since a new group has formed. Proceed with the next new component $\kappa^{(j)} \in \mathcal{G}_j$. Subsequently, proceed with the next network update.

Note that the duration of groups whose activation time is equal to t_0 will be ignored as the actual activation time lies before the beginning of the observation period and is therefore unknown. Likewise, groups which are still active at $t_{\max}$ are not associated with a duration because the actual deactivation times are unknown.

While there exist more sophisticated definitions to define groups and group-durations [5], the definition introduced here is straightforward: A group persists as long as the set of its constituents does not change. The duration of the group is the total period of time this set remains constant.

In order to obtain an overview of the group-size distribution over the whole observation period, the average group-size distribution can be obtained as

$$\overline{n_g} = \frac{1}{T} \int_{t_0}^{t_{\max}} dt \; n_g(t).$$

Note that the inter-contact duration is the time a node spends isolated which is equal to the group life-time of groups of size $g = 1$:

$$\tau_{IC} = \tau_{g=1}.$$

2.3.3 *Empirical Temporal Human Proximity Data Sets*

Recently, two major series of experiments collected temporally resolved human proximity networks, the Copenhagen network study [5, 39, 79] and the SocioPatterns project [40–42]. A brief overview of the experiments and the data sets used in this work is given here.

Copenhagen Network Study

The Copenhagen network study entailed measuring human proximity data, text messages, phone calls, and social media contacts using customized Android smart phones which were handed out to 1,000

students of DTU Copenhagen (Danmarks Tekniske Universitet) [39, 79]. As the focus of this thesis lies on human proximity networks, only this particular data of the whole set is described here, which was measured using the phones' built-in Bluetooth sensors. These sensors were active at all times such that once a phone was within range ($\lesssim 10\,\mathrm{m}$) of another phone, a contact was recorded and collected on a server as a tuple (i,j,t,σ) with i,j being the pair of phones, t denoting the time of initial contact and σ representing the Bluetooth signaling strength. In one of the experiment's studies, it was found that definite thresholds can be set for σ to be representative for two phones being within certain ranges $\lesssim r$ [79]. For the data set analyzed here, the threshold was chosen such that an adjacency matrix entry $A_{ij}(t) = 1$ means that at time t, the phones i and j where within a distance of $< 1-2\,\mathrm{m}$ and an entry $A_{ij}(t) = 0$ in the adjacency matrix implies that at time t, the two phones were located at greater distance. This particular distance was chosen because droplets carrying airborne diseases might typically travel $> 1\,\mathrm{m}$ when breathing indoors or $> 2\,\mathrm{m}$ when coughing indoors [37], distances that are expected to decrease outdoors, hence choosing a mean critical distance of $1-2\,\mathrm{m}$. One might wonder about the effects of under-sampling, since only contacts between those students that carried the custom phone with them were measured. However, in reference [88] it was found that unbiased sampling does not change the statistics of structural and temporal properties of proximity data sets and hence it will be assumed that under-sampling will not cause major deviations when analyzing the data.

The measurement discussed in this thesis started on $t_0 = 0$ which corresponds to Monday, 8th of September 2014, and concluded exactly one week later at $t_{\max} = 168\,\mathrm{h}$. To achieve a temporal discretization, the contacts were binned with a minimal bin size of $\Delta t = 5\,\mathrm{min}$, meaning that when a contact was active at time t, it was set to be active during the whole time bin $[n\Delta t, (n+1)\Delta t)$ with $n\Delta t \leq t \leq (n+1)\Delta t$ and $n \leq 168\,\mathrm{h}/(5\,\mathrm{min}) - 1 = 2015$ being a non-negative integer. Note that the actual measurements on phones were run in a way that an application on the modified smart phones measured the nearby Bluetooth devices and signal strengths for 30 s every 5 min in a desynchronized manner (meaning that the phones did not measure all at the same time). Due to this desynchronized sampling, the data can be considered to be effectively binned over 5 min even though some contacts will have been missed [79]. In total, $N = 412$ individuals participated in the measurement over this period of time. This particular measurement will be referred to as DTU in the following.

Analyzing this contact data one obtains the following properties, visually summarized in Fig. 2.25. At any time of the measurement the network consists of single, densely connected groups (see Fig. 2.25a). The size of these groups varies over time with large groups existing during certain short periods of the day (e.g. lunch time at 12 to 1 pm) and smaller groups dominating during the night; however, isolated

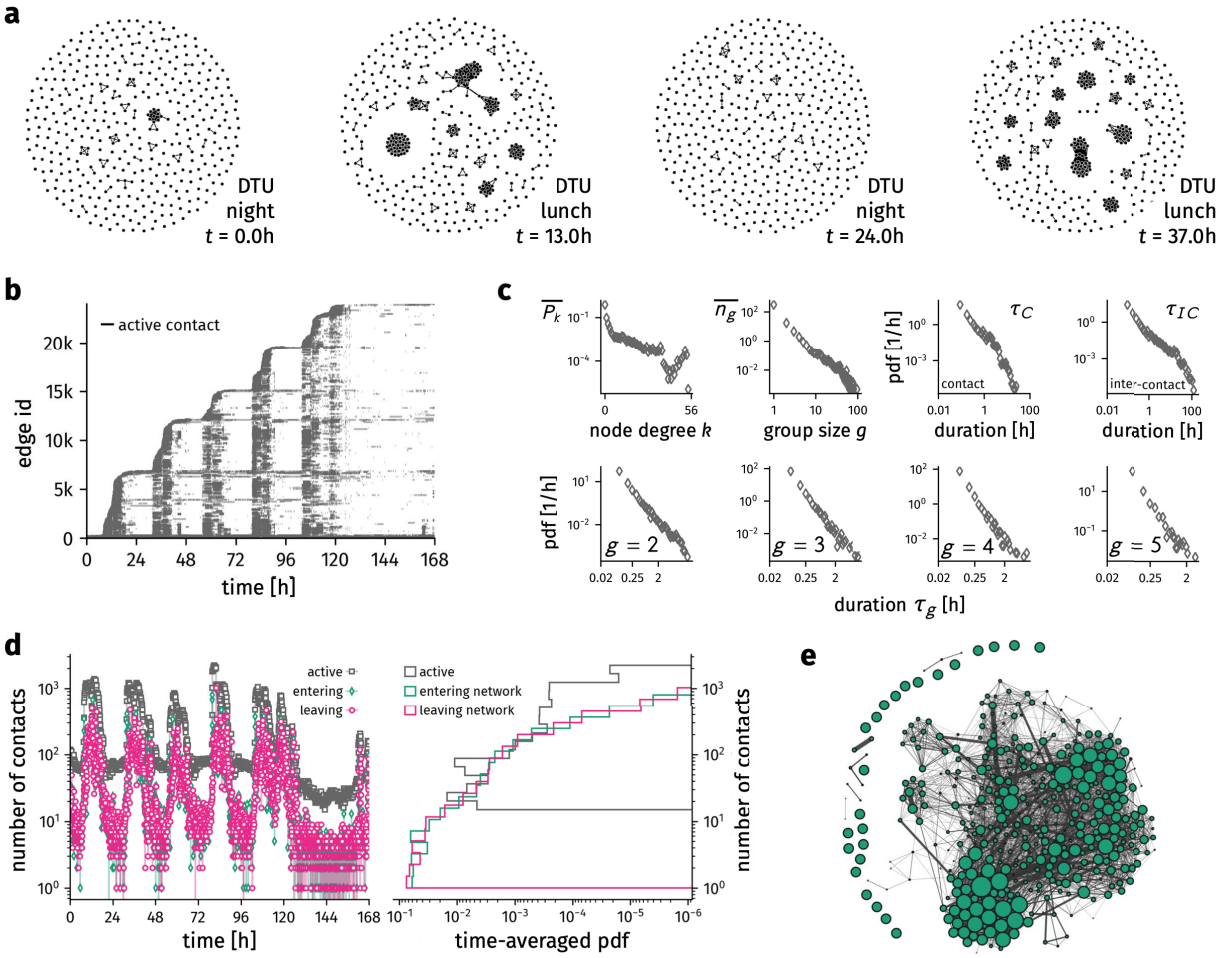


Figure 2.25: The DTU temporal network data set. **(a)** Snapshots at certain times of measurement. At any time, the network consists of densely connected groups, on average larger during the day than during the night. **(b)** A similar activity pattern is shown in an edge activity plot. During the day, more edges are active than during the night. Furthermore, more previously unobserved contacts are measured during the day. **(c)** Almost all structural and temporal observables show heavy-tailed behavior in their time-averaged distributions. An exception is the degree distribution which declines approximately exponentially. **(d)** Active contacts and the number of contacts being (de-)activated follow a circadian pattern but are distributed declining faster than a heavy-tailed distribution. **(e)** The aggregated static network shows 'complex' properties. Note that links associated with the lowest 10% of all link weights are not shown.

nodes are always present. Another way to visualize this behavior is by drawing an *edge activity plot*, in which the x-axis represents time, and discrete values on the y-axis represent edge identifiers (IDs). These IDs are drawn in order of time of first appearence, meaning an edge which appeared earlier than another edge will have a smaller ID. Now, for each edge, a grey bar can be drawn over each interval for which it was active (see Fig. 2.25b). One generally sees that during certain times of the day, previously unobserved contacts are observed for the first time and are replicated during similar times in a cycle of 24 h. During the night, comparably less 'new' contacts are created—however, contacts which are active during the night are more likely to remain active, or switch between their active and inactive states several times. The network is less dense during these times meaning there are less active contacts in the network during the night than there are during the day. This effect is further reflected in Fig. 2.25d which shows the amount of active contacts at time t alongside the number of edges being activated and deactivated in the time interval $[n\Delta t, (n+1)\Delta t)$. These curves follow a clear circadian pattern where during the day there are more edges activated and deactivated per time bin than during the night, while at the same time

there are more edges active in total during the day than there are active during the night. The temporally-averaged distribution of total number of edges, edges being activated per time bin and edges being deactivated per time bin follows a distribution which decays stronger than a heavy-tailed distribution. In Fig. 2.25c, several other time-averaged distributions are shown, focussing first on the degree- and group-size distributions as indicators for the average network structure. The time-averaged degree distribution of the network seems to decline exponentially in two regimes, where for smaller degrees it declines faster than for larger degrees. On the other hand, the group-size distribution has a heavy tail. Temporal descriptors quantifying the network's change are given by the contact-duration, the inter-contact duration and the life-time of groups of size g. One observes that all of these temporal distributions show heavy-tailed behavior, however the contact- and group life-time distributions decline faster than the inter-contact distribution. A possible explanation: keeping a long-lasting face-to-face contact requires more attentiveness than just staying alone and thus might be favorable to some individuals, resulting in longer time periods without contacts than periods of being in contact. Finally, the aggregated network as illustrated in Fig. 2.25e shows properties which qualifies it as being called 'complex', i.e. there are nodes which spent considerably more time being in contact with people than other nodes (heterogeneous node properties) and there is non-zero clustering. The network is furthermore 'sparse' in the sense that the total number of unique contacts observed during the time of measurement is $\ll N(N-1)/2$.

SocioPatterns

For SocioPatterns experiments, face-to-face contact data is measured using radio frequency identification (RFID) chips which are clipped to participants' clothes in the general chest area [40, 88]. The chips are programmed to identify radio signals distributed by other chips when they are nearby ($\lesssim 1-1.5\,\mathrm{m}$), subsequently sending information about those mutual occurences to nearby RFID reading devices with higher detecting range. Since the radio signal from the used chips is shielded by human bodies, the data can approximately be assumed to represent actual face-to-face contact data. For all data sets, the temporal resolution was fixed at $\Delta t = 20\,\mathrm{s}$ using the binning procedure as described above.

The first SocioPatterns data set was collected during the 'Hypertext 2009'-conference where about 75% of all participants volunteered to wear RFID chips on their conference badges, meaning the data set consists of $N = 113$ individuals [42]. The measurement started at $t_0 = 0$, which corresponds to 8am on the 29th of June 2009 and ended at $t_{\max} = 212360\,\mathrm{s} \approx 59\,\mathrm{h}$. This data set will be referred to as HT09 in the following.

The second SocioPatterns data set consists of measured face-to-face contacts of students in a French high school, collected over the

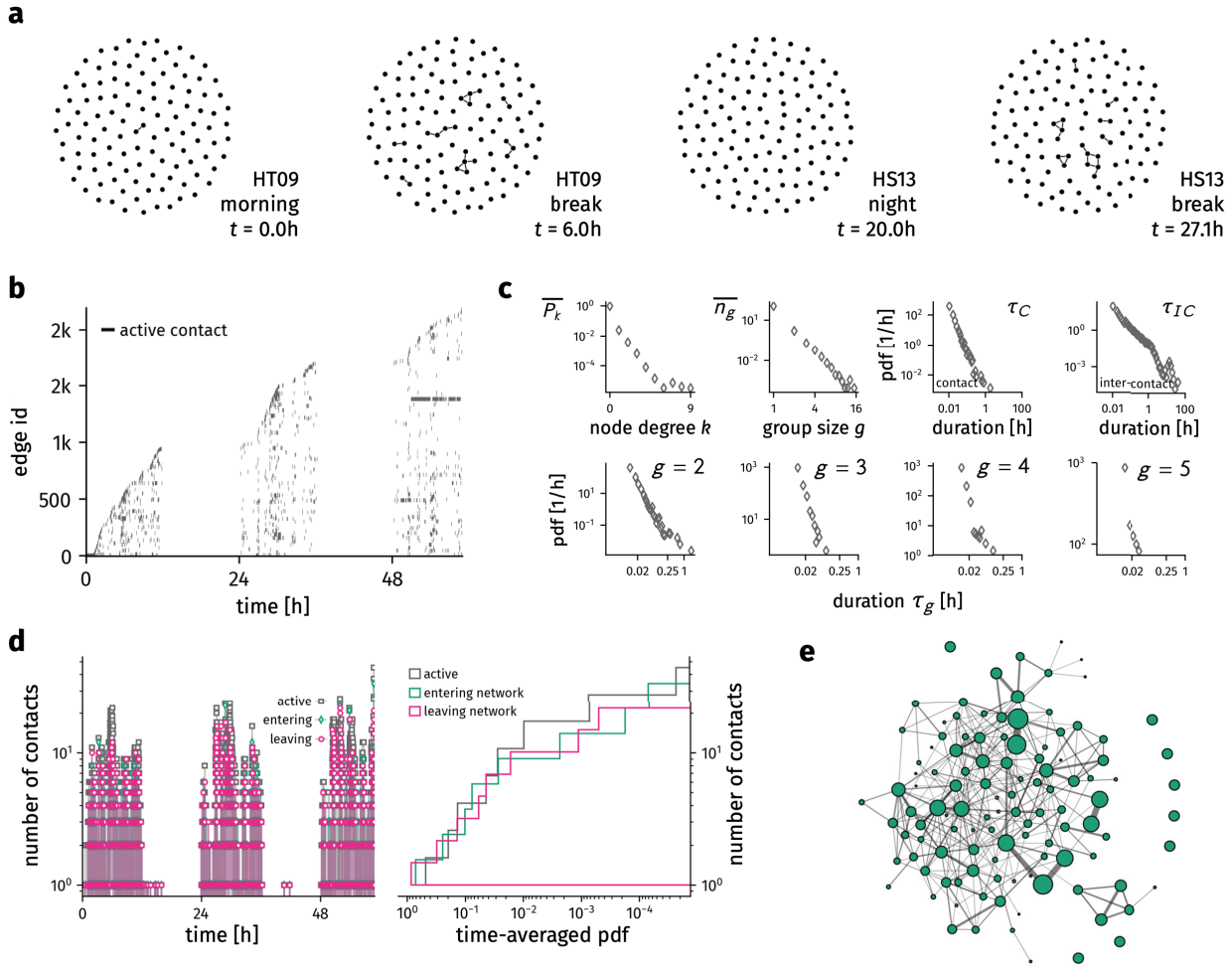


Figure 2.26: The HT09 temporal network data set. **(a)** Snapshots at certain times of measurement. At any time, the network consists of densely connected groups, on average larger during the day than during the night. Generally, groups are much smaller than in the DTU data set. **(b)** An edge activity plot. During the day, more edges are active than during the night. Furthermore, more previously unobserved contacts are measured during the day. **(c)** As for the DTU data, almost all structural and temporal observables show heavy-tailed behavior in their time-averaged distributions. An exception is the degree distribution which declines approximately exponentially. **(d)** Active contacts and the number of contacts being (de-)activated follow a circadian pattern but are distributed declining faster than a heavy-tailed distribution. **(e)** The aggregated static network shows 'complex' properties. Note that links associated with the lowest 10% of all link weights are not shown.

course of five days [41]. About 86% of all students present in this particular school setting participated by wearing RFID chips on lanyards, with the data set consisting of $N = 327$ individuals. The measurement started at $t_0 = 0$, which corresponds to 12pm on the 2nd of December 2013 and ended after $t_{\max} = 363580\,\mathrm{s} \approx 111\,\mathrm{h}$. It will be called HS13 in the following.

The same analyses as for the DTU data set can be done for the SocioPatterns data sets, which are shown in Figs. 2.26 and 2.27, revealing statistical patterns that are qualitatively equal to the ones already described above: The degree distribution decays exponentially, while all other discussed distributions are approximately heavy-tailed. At each time, the network consists of a set of disconnected groups which are densely connected internally. Activity in the network follows a circadian pattern. However, compared to the DTU data, the SocioPatterns data sets show no activity during the night. This effect arises due to the spatial restrictiveness of the SocioPatterns experiments: all contact data can only be recorded in the vicinity of RFID readers on-site (i.e. in the school building for the HS13 data set and in the conference venue for the HT09 data set). The aggregated networks show 'complex' features, i.e. node and link weight heterogeneity as

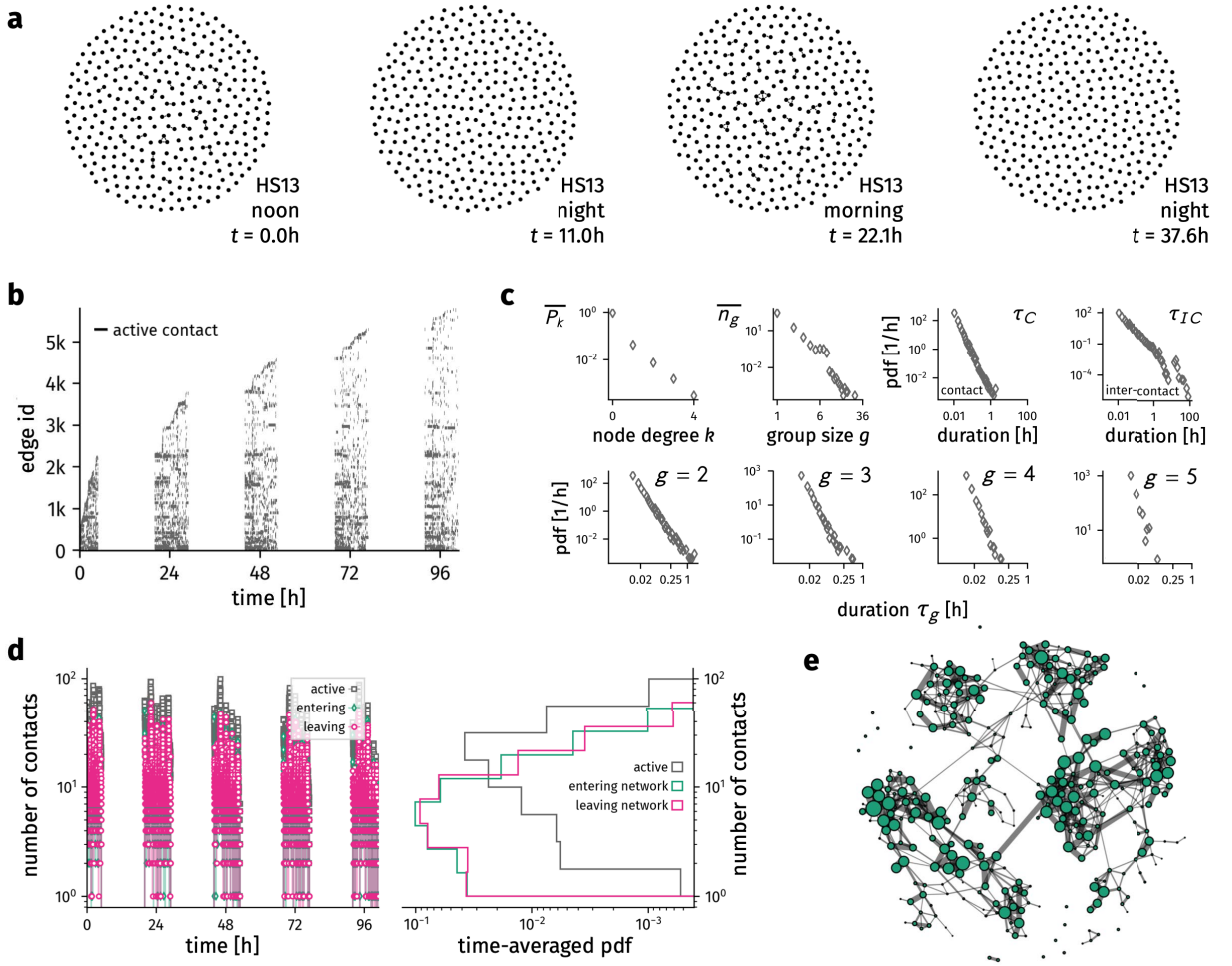


Figure 2.27: The HS13 temporal network data set. Generally, the same properties are observed as in the previously discussed data sets. **(a)** Snapshots at certain times of measurement. **(b)** An edge activity plot. During the day, more edges are active than during the night. Furthermore, more previously unobserved contacts are measured during the day. **(c)** Again, almost all structural and temporal observables show heavy-tailed behavior in their time-averaged distributions. An exception is the degree distribution which declines approximately exponentially. **(d)** Active contacts and the number of contacts being (de-)activated follow a circadian pattern but are not distributed according to a heavy-tailed distribution. **(e)** The aggregated static network shows 'complex' properties including modular clustering. Note that links associated with the lowest 10% of all link weights are not shown.

well as clustering. For the HS13, the aggregated network clearly displays modular clustering (see Fig. 2.27e), too, which is not surprising given that the students spent most of their time in separate classrooms with the same people [41].

Comparing the distributions of structural and temporal observables across data sets reveals that while the shape of the distributions correspond qualitatively (excluding the average degree distribution), the DTU group size and temporal durations are generally greater than the SocioPatterns durations (see Fig. 2.28a and 2.28c). This discrepancy vanishes when one regards the measurement process of both experiments, where contacts are counted in discrete time bins, and acknowledges that the binning time in the DTU experiment is one order of magnitude larger than in the SocioPatterns experiments. Hence, when rebinning the SocioPatterns data to the binning time $\Delta t = 5\,\text{min}$ of the DTU experiment, all distributions besides the degree distribution are of increased correspondence (see Fig. 2.28b and 2.28d). This effect will be discussed in detail in Sec. 7.2.2. All remaining differences are suspected to be caused by the DTU experiment measuring proximity as contacts while the SocioPatterns measures face-to-face contacts. Discrepancies might also be induced by the way in which DTU contacts are measured: phones collect edges for

30 s every 5 min in a desynchronized manner. Considering that each of the experiments was conducted in a different setting with different populations it is indeed surprising that the resulting statistics show similar patterns.

A remarkable result of these analyses is that neither contact, nor inter-contact, nor any group-life time seem to have a typical time-scale as all of them are broadly distributed. In the case of groups, even the group size g does not have an influence on the general shape of its corresponding life-time distribution [43, 88], shown in Fig. 2.28c-d. This is remarkable because it has a direct implication on the outcome of dynamics. In dynamics with definite time-scales like the spread of diseases, one is typically interested in studying how the number of infected changes when infections or recoveries occur 'faster' than the network changes. However, this question would be meaningless considering that there is no definite time-scale to which the disease time-scales can be compared to.

Nonetheless, this lack of a time-scale is not entirely accurate: the cut-off duration of the group-lifetime clearly decreases with increasing group-size, thus introducing at least one time-scale (see insets of Fig. 2.28c and 2.28d). Also, the contact and inter-contact duration distributions show cut-off times which are well below the duration of their corresponding experiments. Furthermore, the distributions are broad but display certain deviations from a perfect power-law. All these observations might imply that the observed distributions are superpositions of several distributions associated with varying time-scales—a hypothesis which is in agreement with the observations of circadian activity in all of the temporal contact data. Former studies, however, explain the emergence of these heavy-tailed distributions by implementing nodes building and cutting contacts with broadly distributed memory kernels, ignoring circadian activity. This will be briefly discussed in the following section.

2.4 Temporal Network Models

In order to investigate both epidemic spreading and the emergence of several properties of temporal networks, several existing modeling approaches are introduced here.

2.4.1 Memory-Kernel Models

Former studies explain the emergence of heavy-tailed distributions in real temporal face-to-face networks on the basis of models in which nodes build and cut contacts using memory kernels [43, 89]. There, it was found that link-specific, link age-specific, and node-age specific activation and deactivation rates are necessary to generate the statistics described in Sec. 2.3.3. Such time-dependent and link-specific rates are associated with a system having memory—in this case the functional dependence of the (de-)activiation rates had to follow a power-law to produce heavy tails in the contact and inter-

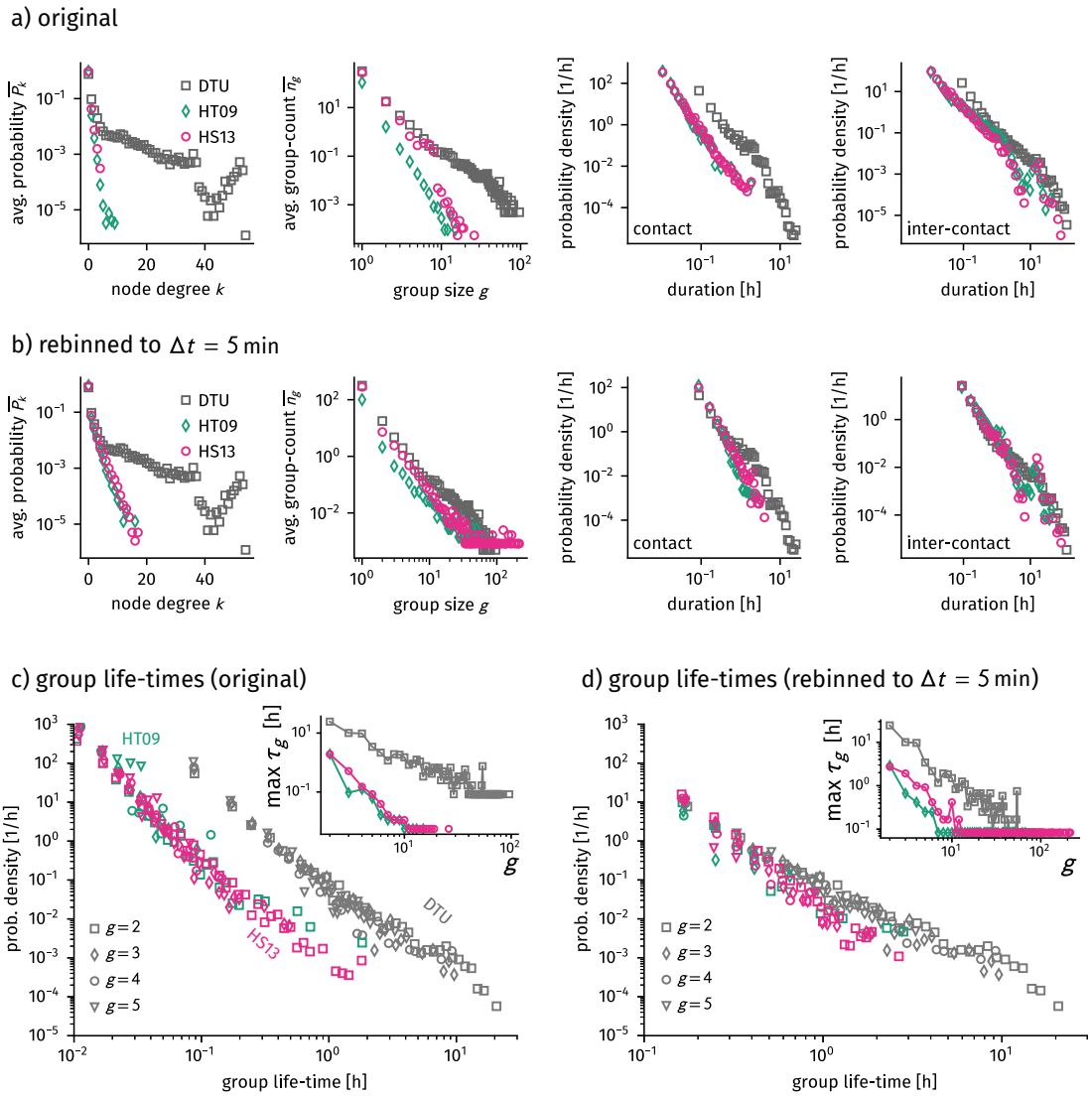


Figure 2.28: Comparison of structural and temporal statistics across DTU, HT09, and HS13 and the influence of binning. **(a)** The observables of the original data sets. **(b)** The observables of the data sets when the data was rebinned to $\Delta t = 5$ min. All distributions but the degree distribution increase correspondence. **(c)** The group-size distributions of the original data sets. Inset: The maximum life-time of groups decreases with size g. **(d)** Rebinning increases the correspondence between the SocioPattern and DTU group-size distributions. Inset: Maximum life-time of groups of size g after rebinning. Rebinning increases correspondence in the scaling of maxτ_g with increasing g.

contact duration distributions. However, the influence of circadian activity was ignored.

One particular model assumes that nodes are always either isolated or part of a larger component in which every node is connected to every other node [43]. In this model, nodes change their connection status with a probability depending on (i) the node's connection status and (ii) the node's last activation time. In discrete time t, each node u is associated with an activation time $t_a^{(u)}$. At every event where the connection status of node u changes, its activation time is updated to $t_a^{(u)} = t$. Denote with $p_g(t, t_a^{(u)})$ the probability of a node u to become active and change its connection status at time t, given that it was last active at time $t_a^{(u)}$ and that it is currently in a group $\kappa(u)$ of size $g = |\kappa(u)|$. The activation probabilities $p_g(t, t_a^{(u)})$

are chosen to follow

$$p_1(t, t_a) = \frac{b_0}{1 + (t - t_a)/N},$$
$$p_{g>1}(t, t_a) = \frac{b_1}{1 + (t - t_a)/N}.$$

At every discrete time step $t \to t + 1$, a node u is randomly chosen and changes its connection status with probability $p_g(t, t_a^{(u)})$. If it does change its connection status, its behavior depends on its group status. If it is isolated, it builds a new pair with a second randomly chosen isolated node v (chosen with probability proportional to $p_{g=1}(t, t_a^{(v)})$). If it is not isolated ($g > 1$), it becomes isolated with probability ι. With probability $1 - \iota$, it invites an isolated node v to its group—again chosen randomly with probability proportional to $p_{g=1}(t, t_a^{(v)})$. All nodes u' involved in this connectivity change update their activation time to $t_a^{(u')} = t$.

Therefore, the model entails basic assumptions about the connection behavior of nodes:

1. Nodes behave differently depending on whether they are connected to other nodes or isolated.

2. A node 'remembers' its activity: the more recently it was active, the more likely it is that it will become active again.

The model is further controlled by three free parameters: the time-independent probability b_0 to become active if a node is isolated, the time-independent probability b_1 to become active if a node is not isolated, and the isolation probability ι.

It was found that in a specific region of the parameter space (b_0, b_1, ι) the temporal network reaches a statistically stable equilibrium state of disjoint groups in which nodes change between being isolated and being in groups of size $g > 1$. The contact, inter-contact, and group durations are shown to follow power-laws for several parameter combinations. While the model therefore reproduces certain features of real-world networks, there are some problems, too:

- The circadially varying activity exhibited in real-world temporal networks is ignored.

- Heterogeneous inter-activity times of single nodes are postulated and do not emerge as a systemic property. This raises the question whether there are some underlying principles following which this heterogeneity arises. In [47, 48] and Sec. 2.1.3 it was shown that Poissonian systems of varying rates may yield inter-event times following heavy-tailed distributions. Thus, heterogeneous inter-event times might be a consequence of circadially varying activity rates.

- The authors do mention the possibility of having a time-independent activity probability p_g which would yield a Poissonian memoryless model, but reject the analysis of such a model due to its unre-

alistic consequences. The exact properties of a similar Poissonian model are therefore still unknown.

- The group structure of the model is postulated. It is therefore unclear whether the phenomenon that nodes only persist in small, densely connected groups might emerge from underlying principles.
- It is ambiguous how to control the exact structural network properties of this model by tuning its three parameters, e.g. to obtain a certain group-size distribution in equilibrium, even though the authors do give an expression for the mean group size. One may wonder whether the introduction of time-varying rates to a Poissonian model eases the evaluations and allows for simpler predictions.
- The model is of qualitative nature in the sense that its parameters do not reflect physical properties of any real-world entity. It is unclear how to choose b_0, b_1, and ι such that the temporal dimension of a resulting model sample would reflect the temporal dimension of an empirically collected temporal network.

To address these problems, a Poissonian group-structured temporal network model will be devised and analyzed in Ch. 7. It will be shown that two rates that describe node behavior are sufficient for group structure to emerge. These rates can be temporally varied to obtain a circadian rhythm in the system. Fixing one of the parameters to set the time-scale, the model therefore has a single parameter which directly controls the group-size distribution. An inference method to map the model parameter to real-world systems is presented.

2.4.2 *Edge Activity Model*

In the simplest continuous-time temporal network model of constant size N, each pair of node (u, v) is associated with two rates [90]. If there is no active link between two nodes, a link is activated with rate ω^+. An active link between two nodes is deactivated with rate ω^-. Hence, the average active duration of a link is $\tau_a = 1/\omega^-$ and the average inactive duration of a link is $\tau_d = 1/\omega^+$. The process of a single node being active or inactive is a telegraph process with the value $A_{uv}(t)$ changing between zero and one with the total rate $\omega = \omega^- + \omega^+$. Therefore, it changes its status with an average period of $\tau = 1/\omega = (\tau_a + \tau_d)/(\tau_a \tau_d)$ and the ergodic probability of a single link existing is given by the proportion of time it is activated such that

$$\begin{aligned} p_{\mathrm{ER}} &= \lim_{T\to\infty} \frac{1}{T} \int_0^T dt\ A_{uv}(t) \\ &= \frac{\tau_a}{\tau} = \frac{\omega}{\omega^-} = \frac{\omega^- + \omega^+}{\omega}. \end{aligned}$$

At each time t, each single link exists with probability p_{ER}. Therefore, at each time t, the network resembles a sample from the Erdős–Rényi (ER) random network model. One may choose the average mean degree $\langle k \rangle$ as the structural control parameter of this model. Any of the three rates ω, ω^-, and ω^+ may then be chosen to control the temporal change of the network, i.e. how fast this local structure changes. Note that while at any point in time the network model is equal to an ER network, its average network will approach a complete network, where each pair of nodes is connected with weight $\overline{A_{uv}} = p_{\mathrm{ER}}$. Due to its similarity to the ER model, this model is sometimes referred to as the dynamic random graph. It is therefore an ideal reference model for real temporal networks because it is both structurally random in the simplest way as well as temporally changing in a simple, random way, based on a Poisson process. In contrast to the original formulation in [90], the continuous-time formulation will be used to sample from this model, with single samples being generated using Gillespie's algorithm (which will be explained in Sec. 3.2.3 based on the detailed formulation given in App. D).

2.4.3 *Other Models*

A larger class of other temporal network models exists which are used to study disease spreading [45, 56, 57, 91]. All of them are defined in discrete time with snapshot length Δt and at each time step, a new configuration is sampled according to some predefined rules. This includes:

- A model where a single configuration is sampled from a set of predefined configurations. Each configuration consists of a number of disjoint completely connected networks of size $g = k + 1$ and isolated nodes [45].
- A model where a single configuration is sampled from a set of predefined configurations. Each configuration consists of a number of disjoint star networks of size $g = k + 1$ and isolated nodes [45]. A star network is a network with a single hub. All nodes other than the hub are only connected to the hub.
- A model where each node u is associated with an activity probability a_u. For each snapshot, each node is activated with probability a_u and connects to k other randomly chosen nodes [46].

The problem of these models is that (a) snapshots adjacent in time are often largely uncorrelated and (b) networks are constant for a time Δt and then abruptly change to a completely different state. This implies that for infitesimal increments in time $dt \ll \Delta t$ a network is either exactly the same as before or completely different. After a few snapshots, one may interpret the networks as well-mixed. In real-world networks, however, the probability of a single link existing clearly follows a smooth function of continuous time. Therefore the 'auto-correlation' of a temporal network should follow a smooth function, as well.

3
Spreading Dynamics

This chapter introduces concepts related to random walks on static networks and simple contagion processes on both static and temporal networks. In Sec. 3.1, theoretical methods concerning the evaluation of the influence of network structure on random walk temporal observables will be introduced. These are the basis for the discussion of random walks on modular hierarchical networks in Ch. 5 and Ch. 6. Parts of Sec. 3.1 are close to texts already published in [92–94]. In Sec. 3.2, the SIS process will be discussed as a simple model of contagion phenomena. This entails the discussion of a widely applied approximation and the circumstances under which this approximation might fail. This section is the basis for the discussion of simple contagion processes on temporal networks in Ch. 8.

3.1 Discrete-Time Random Walks

Random walks have been studied extensively for more than a century and emerged as an efficient descriptive model for spreading and diffusion processes in physics, chemistry, economics, biology, social sciences, epidemiology, and computer science [9, 32–34, 95–97]. Because of their wide applicability and relevance to dynamic phenomena, random walk processes have become a topic of interest particularly for analyses of dynamics on complex networks [96]. The calculation of a resistor network's effective resistance [10], synchronization phenomena in networks of coupled oscillators [9, pp. 138–141], the global spread of infectious diseases on the global air traffic network [98, 3, 4], and ranking the importance of single websites in the world wide web [2] are just a few examples of systems that have been investigated based on concepts derived from random walk theory.

The understanding of temporal aspects of stochastic processes and how different network structures influence their equilibration process is of special importance to make logistic decisions or design structures which facilitate or mitigate fast diffusion. Consequently, theoretical work focused on understanding the connection between network structure and relaxation time scales or first passage times (FPTs), the time it takes a single walker to travel from one node to another [10, 99–102]. Often, however, predictions rely on effec-

tive or average medium approximations where instead of the actual network structure, connection probabilities are considered [103–105]. While this was shown to be an appropriate approximation for dense networks, strong deviations may arise for networks of small mean degree [103]. Other work developed a heuristic to explain mean first passage times on the basis of structural properties of specific network samples, however, relies on the approximation of the network to be locally tree-like, i.e. to possess a vanishing clustering coeffcient [100]. This heuristic will be explained in Sec. 3.1.5. As demonstrated in the last chapter, both the assumption of a large mean degree as well as vanishing clustering are not met in e.g. social networks. Therefore, a new heuristic will be developed in Sec. 5.2.2 which explains qualitative features of the mean first passage time on the basis of the clustering coefficient and the degree distribution.

Another important time scale is captured by the cover time of a random walk process which is the time needed to visit *every* node in a network, introduced in Sec. 3.1.6. This quantity has important practical applications from biology to computer science, for instance, for estimating how long it will take to distribute a chemical or certain commodity to every node in a network, as a measure for navigability in multilayer transportation networks [106], or for exhaustive search processes [107]. In Ch. 6, a new heuristic is introduced to compute the mean cover time with high precision on arbitrary networks. To the best knowledge of the author it is the first theoretical method to explictly connect first passage time statistics to the cover time.

Note that while most of the research concerning random walks focuses on diffusion in continuous media [33], this thesis solely concentrates on random walks in discrete media, i.e. networks. Therefore, a discussion of random walk theory in continuous media will be omitted.

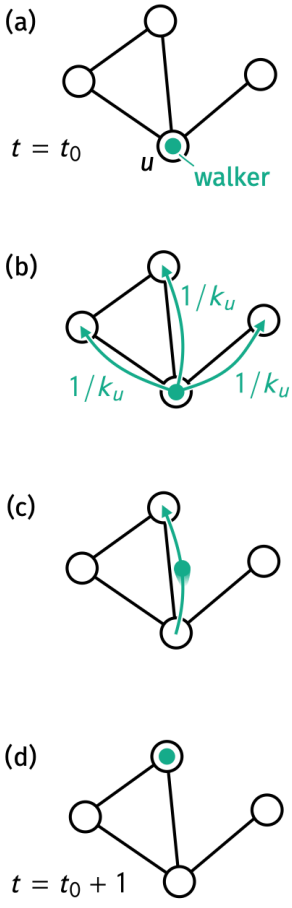


Figure 3.1: Illustration of a single time step in a random walk process on an undirected unweighted network. **(a)** A walker resides on a node u at $t = t_0$. **(b)** Each neighbor $v \in \mathrm{Nei}(u)$ has equal probability of being chosen as the walker's target for this step. Since the probability that any of them is chosen has to sum to one, each of them is chosen with probability $1/k_u$. **(c)** The walker chooses one of the neighbors at random to jump to. **(d)** The walker now resides on the chosen node at time $t = t_0 + 1$.

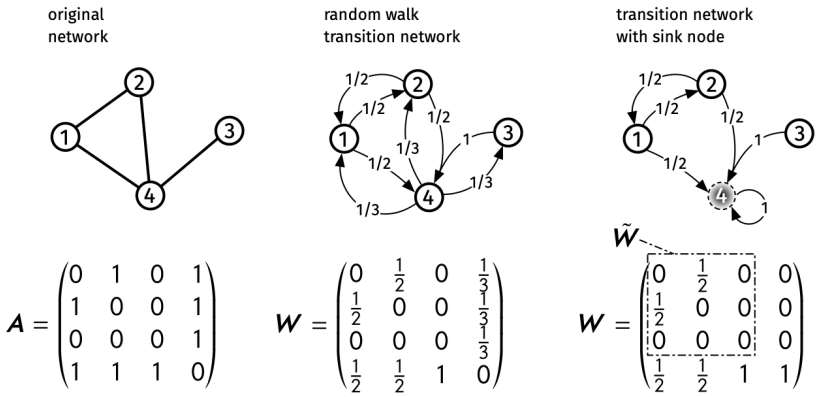
3.1.1 *Random Walks on Networks*

A simple discrete-time random walk process on an undirected network without memory is defined as follows. At time point $t \in \mathbb{Z}$, a random walker resides on a node u with degree k_u. It then chooses one of u's neighbors, each uniform at random with probability $1/k_u$. Subsequently it jumps to the neighbor v to reside there at time $t+1$. This process is then repeated indefinitely. It is schematically illustrated in Fig. 3.1.

The description above implies that this random walk process is a Markov process with transition matrix $\boldsymbol{W}$. Given that a random walker resides on node u, the probability for the jump $u \to v$ is given as the probability that u and v are connected and the probability that v is chosen from all neighbors such that

$$W_{vu} = A_{vu} \times \frac{1}{\sum_{j=1}^{N} A_{ju}} = \frac{A_{vu}}{k_u}, \tag{3.1}$$

as illustrated in Fig. 3.2. Hence the probability $p_v(t+1)$ for a walker to be on node v at time $t+1$ is given as the probability $p_u(t)$ to be



$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \qquad W = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{3} \\ \frac{1}{2} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & 1 & 0 \end{pmatrix} \qquad W = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 & 1 \end{pmatrix}$$

Figure 3.2: Illustration of the random walk transition matrix as a weighted directed network and the transition matrix with a sink node. **(Left)** Given the original adjacency matrix with entries A_{vu}, **(center)** the corresponding entry W_{vu} in the random walk transition matrix is computed as the probability to transition from u to v which is given as Eq. (3.1). **(Right)** Introducing a sink at node 4, a random walker cannot leave anymore once it arrived there, represented by the column $(0, 0, 0, 1)^T$ in the matrix and by a self-loop in the illustration. The reduced transition matrix $\tilde{W}$ is constructed by dropping both row and column associated with node 4.

on *any* node u at time t and the probability to transition from u to v. The whole process is thus governed by the master equation

$$p_v(t+1) = \sum_{u=1}^{N} W_{vu} p_u(t). \tag{3.2}$$

In equilibrium, walker probability has to stay constant on each node meaning that the total walker influx of a node has to equal the total walker outflux. Hence, the equilibrium walker density $p_u^\star$ has to meet

$$\underbrace{\sum_{v \in \text{Nei}(u)} \frac{1}{k_u} p_u^\star}_{=\text{outflux}} = \underbrace{\sum_{v \in \text{Nei}(u)} \frac{1}{k_v} p_v^\star}_{=\text{influx}}.$$

This equation is satisfied when $p_u^\star = 0$ or $p_u^\star \propto k_u$. The first solution is only valid for initial conditions $p_u(t=0) = 0$ for all u since $\boldsymbol{W}$ is a walker-conserving operator. Hence, the non-trivial solution with non-zero initial condition is

$$p_u^\star = \frac{k_u}{\sum_{v=1}^{N} k_v} = \frac{k_u}{2m} = \frac{k_u}{N \langle k \rangle} \tag{3.3}$$

where $\sum_{v=1}^{N} k_v = 2m = N \langle k \rangle$ was used to obtain the total number of edges in the network m.

3.1.2 *Graph Laplacian*

Another way to find the equilibrium density is to introduce the so-called unnormalized graph Laplacian. In a matrix-vector notation, the entries p_u of the column vector $\boldsymbol{p}$ represent the random walker density at each node u. Let $\boldsymbol{D}$ be a diagonal matrix where each diagonal entry is $D_{uu} = k_u$ and each off-diagonal entry is $D_{uv} = 0$. Then the transition matrix is given as

$$\boldsymbol{W} = \boldsymbol{A}\boldsymbol{D}^{-1}$$

and the master equation can be written as

$$\boldsymbol{p}(t+1) = \boldsymbol{W}\boldsymbol{p}(t) = \boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{p}(t) \tag{3.4}$$

such that the equilibrium density is determined by

$$\boldsymbol{p}^\star = \boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{p}^\star.$$

Rewriting this equation gives

$$\begin{aligned} 0 &= \left(\mathbb{1} - \boldsymbol{A}\boldsymbol{D}^{-1}\right) \boldsymbol{p}^\star \\ &= \left(\boldsymbol{D} - \boldsymbol{A}\right) \boldsymbol{D}^{-1} \boldsymbol{p}^\star \\ &= \boldsymbol{L}\boldsymbol{D}^{-1} \boldsymbol{p}^\star. \end{aligned}$$

Introducing the unnormalized graph Laplacian

$$L_{vu} = \delta_{vu} k_v - A_{vu}, \tag{3.5}$$

one recognizes that $\boldsymbol{D}^{-1}\boldsymbol{p}^\star$ is an eigenvector of this matrix $\boldsymbol{L}$ with corresponding eigenvalue 0. Furthermore, it has to be proportional to the column vector $\mathbf{1} = (1, 1, \ldots, 1)^T$ since

$$\sum_{u=1}^{N} L_{vu} = k_v \sum_{u=1}^{N} \delta_{vu} - \sum_{u=1}^{N} A_{vu} = 0.$$

Thus, the equilibrium density is given as

$$\boldsymbol{p}^\star \propto \boldsymbol{D}\mathbf{1}$$

such that after normalization one obtains Eq. (3.3).

The graph Laplacian obtained its name from the fact that it arises as a discrete generalization of the continuous Laplacian operator in a number of contexts where linear coupling plays a role. Examples are coupled heat baths where it generalizes the heat equation [108], coupled oscillators where it plays the continuous Laplacian's role in the wave equation [9, pp. 138 ff.], and in coupled resistors where it arises from applying Kirchoff's laws to the meeting points of resistors [10].

As will be shown later, its spectral properties play important roles for all the dynamics described above as well as explicitly for the computation of passage times in simple random walks. The graph Laplacian's eigenvalues are strictly non-negative and ordered as

$$0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N.$$

It is, e.g. possible to show that the multiplicity of the eigenvalue $\lambda_1 = 0$ is equal to the number of disconnected components of the network meaning that in a single-component network one strictly finds $0 = \lambda_1 < \lambda_2$. Furthermore, the graph Laplacian's eigenvectors

$$\boldsymbol{\mu}_i = (\mu_{i1}, \mu_{i2}, ..., \mu_{iN})^T \tag{3.6}$$

will be referenced in this work where $\boldsymbol{\mu}_i$ is the column eigenvector corresponding to eigenvalue λ_i. As noted above, the vector associated with the smallest eigenvalue $\lambda_1 = 0$ is $\boldsymbol{\mu}_1 = (1, 1, \ldots, 1)^T$.

3.1.3 Mixing Time

This section introduces the *mixing time*, a scale measuring how fast a simple random walk process approaches the equilibrium distribution. The random walk master equation in the form of Eq. (3.4)

contains the matrix $\boldsymbol{W}$ which, as a stochastic matrix, has the property that its largest eigenvalue is always $\omega_N = 1$ if the eigenvalues are ordered as $\omega_1 \leq \cdots \leq \omega_{N-1} < \omega_N$. It was already shown in Sec. 3.1.1 that the corresponding eigenvector $\boldsymbol{w}_N$ is proportional to Eq. (3.3) as $\boldsymbol{p}^\star = \alpha_N \boldsymbol{w}_N$. Each initial condition $\boldsymbol{p}(0)$ can thus be expressed as a linear combination of $\boldsymbol{W}$'s orthonormal eigenvectors $\boldsymbol{w}_i$ as $\boldsymbol{p}(0) = \sum_{i=1}^{N} \alpha_i \boldsymbol{w}_i$ such that

$$\boldsymbol{p}(t) = \boldsymbol{W}^t \boldsymbol{p}(0) = \boldsymbol{p}^\star + \sum_{i=1}^{N-1} \omega_i^t \alpha_i \boldsymbol{w}_i.$$

Since all eigenvalues are sorted one can see that the leading term in this equation is proportional to ω_{N-1}^t and that with increasing t, all non-equilibrium contributions vanish. In fact, it is possible to show that the χ^2-distance between the time-dependent state and the equilibrium state is reduced by a constant factor approximately every

$$t_{\text{mix}} = \frac{1}{1 - \omega_{N-1}} \tag{3.7}$$

time steps, which is the time scale being referred to as *mixing time* in this work [109].

3.1.4 (Mean) First Passage Time

Central questions for random walks are often connected to first passage times (FPTs), e.g. the mean first passage time τ_{vu} (MFPT) between two nodes u and v, which is the mean number of time steps it takes a random walker to arrive at node v if it started at node u. The following derivations closely follow the argumentation in [32, pp. 147 ff.]. It is presented here because the derivations to find the mean first passage time on averaged modular hierarchical networks will be related (see Sec. 5.1).

The MFPT cdf on any node can be found by introducing a deep trap (or 'sink') at the target node v, which refers to a situation where once a random walker hits the node v, it will stay there forever (see Fig. 3.2, right panel). Mathematically, this situation can be described by dropping the v-th row and column in the transition matrix $\boldsymbol{W}$ to obtain the adjusted transition matrix $\tilde{\boldsymbol{W}}$ as well as the v-th entry in the corresponding vectors. The adjusted master equation is then

$$\tilde{\boldsymbol{p}}(t+1) = \tilde{\boldsymbol{W}} \tilde{\boldsymbol{p}}(t) = \tilde{\boldsymbol{W}}^{t+1} \tilde{\boldsymbol{p}}(0).$$

Since all walker probability can flow into the deep trap but not out, the probability to find a random walker being trapped in node v at time t is

$$p_v(t) = 1 - \sum_{u \neq v} p_u(t) = 1 - \tilde{\mathbf{1}}^T \tilde{\boldsymbol{p}}(t).$$

Here, $\tilde{\mathbf{1}} = (1, 1, \ldots, 1)^T$ is an $(N-1)$-dimensional column vector filled with ones. The probability to hit v first at time t is then given as $p_v(t) - p_v(t-1)$ such that the mean first passage time $\tilde{\tau}$ is given

as

$$\begin{aligned}\tilde{\tau} &= \sum_{t=0}^{\infty} t\big[p_v(t) - p_v(t-1)\big] \\ &= \tilde{\mathbf{1}}^T \tilde{\boldsymbol{D}} \tilde{\boldsymbol{L}}^{-1} \tilde{\boldsymbol{p}}(0).\end{aligned}$$

Hence, any mean first passage times are directly related to the graph Laplacian. However, instead of working with the inverse of the reduced Laplacian, it has been shown [110] that the MFPT from start node u to end node v can be found using the graph Laplacian's spectrum as

$$\tau_{vu} = \frac{1}{2} \sum_{i=1}^{N} k_i \left[R_{vu} + R_{vi} - R_{ui}\right] \tag{3.8}$$

where [10]

$$R_{vu} = \sum_{i=2}^{N} \frac{1}{\lambda_i} \left|\mu_{iv} - \mu_{iu}\right|^2 \tag{3.9}$$

with μ_{iv} being the v-th entry of the graph Laplacian's i-th eigenvector as per Eq. (3.6). Hence, finding the mean first passage time between any two nodes reduces to finding the diagonalization of the unnormalized graph Laplacian.

3.1.5 Global Mean- and Pair-Averaged First Passage Time

Another central time observable in the theory of random walks is the *global mean first passage time* (GMFPT) which can be used as a measure of centrality for node v: A node that is quickly reachable from anywhere in the network might be interpreted to be "important". Combining Eq. (3.8) and Eq. (3.9) and noticing that the GMFPT is given as the average over all start nodes yields

$$\tau_v = \frac{1}{N-1} \sum_{u \neq v} \tau_{vu} \tag{3.10}$$

$$= \frac{N}{N-1} \sum_{i=2}^{N} \frac{1}{\lambda_i} \left(2m\mu_{iv}^2 - \mu_{iv} \sum_n k_n \mu_{in}\right), \tag{3.11}$$

a result first derived in [101].

While the result is exact, it is not particulary informative about how structural features of a network contribute to the value of the GMFPT. To resolve this, a heuristic result was derived by Lau *et al*. under the assumption that a simple random walk equilibrates quickly on a network, meaning that the mixing time is $t_{\text{mix}} \ll N$ [100]. Their argumentation is outlined below in greater detail as it will provide the basis for a generalization presented in Sec. 5.2.2.

The following is close to the text published in [93]. The approach by Lau *et al.* treats the time dependent walker concentration $p_u(t)$ on any node u at time t as separated distributions

$$p_u(t) \approx p_u p_{\text{total}}(t)$$

accounting for a time-independent distribution p_u of walker density on node u and the total density of walkers spread throughout

the network $p_{\text{total}}(t)$. In this scenario, walkers are spread over the whole network quickly, reaching equilibrium density. Then, the total amount of walkers steadily decreases through the sink node, almost equally on all nodes. The total amount of walker density leaving the network to end up in sink v is proportional to the amount of walkers leaving neighbors of i. The outflux of walker density at time step $t \to t+1$ is hence

$$\beta_v(t) = \sum_{u \in \text{Nei}(v)} \frac{p_u(t)}{k_u},$$

meaning that the total walker density in the network is changing as

$$\begin{aligned} p_{\text{total}}(t+1) - p_{\text{total}}(t) = \beta_v(t) &= p_{\text{total}}(t) \sum_{u \in \text{Nei}(v)} \frac{p_u}{k_u} \\ &= \beta_v p_{\text{total}}(t) \end{aligned}$$

such that, by recursion,

$$p_{\text{total}}(t) = p_{\text{total}}(0)(1-\beta_v)^t \overset{\beta_v \ll 1}{\approx} p_{\text{total}}(0)\exp(-\beta_v t),$$

where, with the prefactor $p_{\text{total}}(0) = 1$, this coincidentally yields the cumulative distribution function (cdf) of the first passage time at target node i as

$$p_v(t) = 1 - \exp(\beta_v t).$$

The decay rate is given as

$$\beta_v = \sum_{u \in \text{Nei}(v)} \frac{p_u}{k_u}. \tag{3.12}$$

In order to compute the decay rate β_v and hence the GMFPT as $\tau_v = \beta_v^{-1}$, it is necessary to estimate the sink's neighbors' walker concentration p_u. Lau *et al.* propose that second neighbors i of sink node v with distance $d(v,i) = 2$ are approximated to carry equilibrium density $p_i = p_i^\star = k_i / N\langle k \rangle$. Doing so gives

$$p_u = \sum_{i \in \text{Nei}(u) \setminus v} \frac{p_i}{k_i} = \frac{1}{N\langle k \rangle}(k_u - 1). \tag{3.13}$$

This includes the assumption that all walkers flowing into neighbor node u originate from second neighbors i implying that neighbors of v are not connected and hence v has local clustering coefficient $C_v = 0$. Thus, the decay rate of sink node v is given as

$$\beta_v = \frac{1}{N\langle k \rangle} \sum_{u \in \text{Nei}(v)} \frac{k_u - 1}{k_u}.$$

Replacing all k_v neighbors u with average neighbors and using $\left\langle k_{\text{neigh}}^{-1} \right\rangle = \langle k \rangle^{-1}$ which is only valid for sufficiently random networks [32, pp. 377–379] yields

$$\beta_v = \frac{k_v}{N\langle k \rangle}\left(1 - \frac{1}{\langle k \rangle}\right).$$

Hence, in this approach, the heuristically estimated GMFPT would be given as

$$\tau_v^{\text{es}} \geq \frac{N \langle k \rangle}{k_v} \frac{1}{1 - \langle k \rangle^{-1}}. \tag{3.14}$$

It is merely a lower bound since the time steps it takes for the process to equilibrate are ignored. It was shown that the approximation of the exact GMFPT with this lower bound holds within small relative errors for Erdős–Rényi (ER) and Barabási–Albert (BA) networks, as well as for a variety of real-world networks [100]. One of the main contributions of this work will be the extension of this result to networks with non-zero clustering coefficient (Ch. 5).

Based on the GMFPT, the *pair-averaged first passage time* measures the average search time between any pair of nodes as

$$\langle \tau \rangle = \frac{1}{N} \sum_{v=1}^{N} \tau_v. \tag{3.15}$$

An optimal search will take the shortest path between two nodes such that its run time for a network can be measured by the average shortest path length. The pair-averaged first passage time is the equivalent of a random search process. It therefore poses an upper bound for the average search time of any informed search.

3.1.6 Cover Time

Another temporal characteristic with practical relevance is the mean cover time T_u, defined as the mean number of steps it takes a random walker starting at node u to visit every other node at least once. For various network models simple heuristics concerning the asymptotic scaling of the mean cover time as a function of network size N have been derived [99, 111, 112]. For ER, BA and fully connected networks it was shown that

$$\langle T \rangle \simeq \alpha \, N \log N,$$

with model-specific prefactor α. Here, $A_N \simeq B_N$ is chosen to denote $\lim_{N\to\infty} A_N / B_N = 1$. Such scaling relationships are useful for comparative analyses, e.g. when networks for different sizes of the same class are compared but are less helpful when actual expected cover times need to be computed for empirical networks where N is fixed. Unfortunately, a general procedure for estimating the actual cover time for arbitrary complex networks, as well as the connection between the mean cover time and FPT observables does not seem to exist. One of the contributions of this work will be the introduction of a simple heuristic to estimate the cover time with high precision for any network, see Ch. 6.[1]

[1] Parts of this subsection are similar to the text published in [92], © 2017 American Physical Society.

3.2 The SIS Contagion Model

Random walkers do not replicate. Similarly, they do not cease to exist. In absence of a sink, the total amount of random walkers in a network will remain constant. Therefore, random walks are a

convenient model for processes like sending commodities through networks. However, they lose their applicability when information is able to replicate and to vanish, as it does during the spread of a disease or of gossip.

In epidemiological contexts, researchers model the dynamics of infectious diseases in order to make accurate predictions about the expected time-line of large-scale epidemics, or to understand certain patterns in temporal variations of the number of cases. To this end, compartmental models proved to be a valuable tool [54, 55, 113]. Within these models, people are assumed to be in certain states (compartments), e.g. susceptible for an infection (S), infected and infectious towards others (I), or recovered and immune (R). Following defined stochastic rules, they change between these compartments. Based on models like this, one may explain the rise of measles in pre-vaccine New York City [54, pp. 8], the number of cases of influenza over time in a boarding school [53], or how diseases spread through air traffic [3, 4]. In one of the simplest abstractions of epidemic spreading, people are either susceptible or infected—upon contact, a susceptible person can contract the disease from an infected one. Additionally, an infected person may recover to become susceptible again. This dynamic is called the susceptible-infected-susceptible (SIS) compartmental model and is one of the backbones of computational epidemiology. Central phenomenological questions which are approached by means of the SIS model regard the prediction of the size of the infected cluster in the endemic state (number of infected at a certain time) and measures to eradicate the disease. These may include the reduction of contacts by quarantine or vaccination, or the reduction of the mean time a node is infectious using medication.

In this section, the SIS-model is introduced from first principles alongside its various formulations as a system of integrative equations. The purpose of studying the SIS-model in this thesis is neither to study a specific disease nor to make accurate predictions for real-world situations. Rather, the basic model is studied in detailed to reveal certain misconceptions about approximations in its application to time-varying networks.

3.2.1 Mean-Field SIS and the Epidemic Threshold

For an SIS process evolving in a system of N nodes, two events may occur at any time, following the reaction equations

$$S + I \xrightarrow{\eta} 2I$$
$$I \xrightarrow{\rho} S.$$

In these equations, S denotes a single susceptible individual, while I denotes a single infected individual. The first equation describes the disease transmission from an infected to a susceptible person upon a contact with infection rate η. The probability for a single SI-contact to produce an infected individual within an infinitesimal step of length dt is denoted as $\tilde{\eta} = \eta\, dt$. The latter reaction describes an

event where an infected spontaneously recovers with probability $\tilde{\rho} = \rho\, dt$. Now let S denote the total number of susceptibles and $\langle s \rangle = S/N$ the probability for a randomly chosen node to be healthy. Likewise, the probability for a randomly chosen node to be infected is $\langle i \rangle = I/N = 1 - \langle s \rangle$ because $S + I = N$ at all times. A simple way to consider contact structure is to assume a well-mixed system, which here will be interpreted as every node either having contact with $n \gg 1$ other nodes of the system, *or* an infected node changing its contacts very fast with respect to the time-scale of the disease (i.e. very often before it recovers). In this case, a node encounters a subpopulation representative of the total population.[2] The respective equation[3] for the evolution of the probability to be healthy is then given by

$$\langle s(t + dt) \rangle = \langle i(t) \rangle \times \tilde{\rho} + \langle s(t) \rangle \left(1 - \tilde{\eta}[n \times \langle i(t) \rangle]\right). \tag{3.16}$$

This equation should be read as follows. The probability to be healthy after the next infinitesimal increment of time is given by the probability to have been infected before and recover *or* to have been healthy before and *not* being infected by any of the n infected nodes it is in contact with. Here it is assumed that each of those neighbors is infected with uniform probability $\langle i(t) \rangle$.[4] Therefore, the probability for any node to be susceptible evolves in the continuous-time limit $dt \to 0$ according to

$$\frac{d \langle s \rangle}{dt} = -\eta n \langle s \rangle (1 - \langle s \rangle) + \rho(1 - \langle s \rangle).$$

This equation does not have a closed-form solution to obtain the change of susceptibles over time. However, the steady state can be found, which describes the probability to be healthy after a substantial amount of time has passed and can be determined by solving $d \langle s \rangle / dt = 0$ for $\langle s \rangle$, yielding the solution

$$\langle s^\star \rangle = \frac{\rho}{\eta n}. \tag{3.17}$$

Given that each node is connected to n other nodes at all times, one may introduce a dimensionless parameter, the basic reproduction number

$$R_0 = \frac{\eta n}{\rho}. \tag{3.18}$$

This quantity gives the average number of individuals a single infected individual will infect in an otherwise healthy population before it recovers. Applying this to Eq. (3.17) and using $\langle i^\star \rangle + \langle s^\star \rangle = 1$, the epidemic response curve as a function of the basic reproduction number is given as

$$\langle i^\star \rangle = 1 - \frac{1}{R_0}, \tag{3.19}$$

shown in Fig. 3.3. This result is important because it shows that a well-mixed system, on which an SIS process spreads, undergoes a phase transition with R_0 being the control parameter. This parameter can be interpreted as the average number of secondary infections

[2] This is an approximation which originated in the study of chemical systems where a single chemical unit like a molecule can be considered to be part of a subvolume that is representative of the whole solution, see e.g. [73].

[3] It really should rather say "approximate equation" for reasons which will be explained later on.

[4] Note that these equations further assume that only a single event per infinitesimal time step may occur as is the basis for a Poisson process like this (see Sec. D.1.1)—therefore the *or* operations leading to $n \times \langle i(t) \rangle = \sum_{i=1}^{n} \langle i(t) \rangle$ are *exclusive or* operations, otherwise they would have to be *and* operations of the probabilities not to be infected (this will be further discussed below).

caused by a representative node that is initially infected in an otherwise fully susceptible population. In the case of $R_0 < 1$, the initially infected node causes, on average, less than one secondary infection—therefore, the disease is not sustainable and dies out. For $R_0 > 1$ it triggers a cascade of infection events and as such the disease becomes *endemic*. Hence, this critical value of $R_{0,c} = 1$ is usually referred to as the *epidemic threshold*. Working solely in this picture, epidemiology can be viewed as the endeavor to find ways to either reduce the infection rate η or increase the recovery rate ρ using, e.g., medical interventions until $R_0 \leq 1$ such that the disease is eradicated. Measures which can be undertaken to reduce R_0 with structural change are vaccinations, which renders susceptible nodes to be un-infectable, effectively reducing the number of contacts n until the disease dies out. Another measure would entail taking an infected quarantine which similarly reduces the subpopulation in contact with an infected individual.

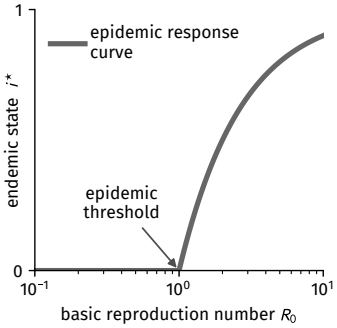


Figure 3.3: The epidemic response curve Eq. (3.19) represents the fraction of infected people in equilibrium, functionally dependent on the basic reproduction number R_0 as given by Eq. (3.18). The value $R_{0,c}$ at which the endemic state $\langle i^\star \rangle$ first vanishes is called the epidemic threshold. Above the epidemic threshold, two equilibrium points exist: the unstable disease-free state $\langle i^\star \rangle = 0$ and the stable endemic state $\langle i^\star \rangle > 0$. Below the epidemic threshold, only the disease-free state exists.

Of course, the assumption that a single average person is in permanent contact with either $n \gg 1$ individuals at all times or changes its n contacts quickly to certain other random individuals before an infected node recovers is counter-intuitive. As has been shown in Sec. 2.3.3, in reality people have

1. heterogeneous contact activity (reflected by a broad degree distribution in the aggregated network),
2. circadially varying amounts of contacts,
3. contact with a smaller amount of people as $\langle k \rangle \equiv n \sim \mathcal{O}(1)$, and
4. are mostly in contact with a recurring set of other people (instead of random mixing).

As such, all assumptions to make the mean-field approximation valid are violated in real systems. In the following, the implications of these violations will be discussed.

3.2.2 *Static Networks and the Individual-Based Markov-Chain Approximation*

In a first abstraction, the loosely defined "contact" structure of the mean-field approach is replaced by a quenched representation of the system, i.e. a network of N nodes with adjacency matrix $\boldsymbol{A}$. In case of long durations of infection, this network might represent the aggregated or averaged contact network between people. Second, it is assumed that a node u can be in binary states where $s_u = 1$ if node u is susceptible and $s_u = 0$ if it is not, which may change as a step function in continuuous time (note that the infectious state is given as $i_u = 1 - s_u$). Therefore, the system is in one of 2^N possible states at any time with constant transition rates between those states. Hence, the process can be described as a Markov process. In the widely applied individual-based Markov-chain approximation [45, 49, 50, 57,

91, 114], the original Markov process is approximated by the transition of single node states. What follows is roughly based on the derivations presented in [32, pp. 667 ff.] and [115, pp. 462 ff.].

Similar to the argumentation for Eq. (3.16), the expectation value of node u's health state evolves as

$$\langle s_u(t+dt)\rangle = \left\langle [1-s_u(t)]\rho dt + s_u(t)\left[1-\eta\sum_{v=1}^{N} A_{uv}(1-s_v(t))\right]dt \right\rangle .$$

Here, the expectation value is taken over the ensemble of possible states the system might be in at time t if it was started with the same initial conditions. Since s_u is a Bernoulli random variable, $\langle s_u(t)\rangle$ denotes the probability of node u to be healthy at time t [115, p. 462]. The corresponding ODE is given by

$$\frac{d\langle s_u\rangle}{dt} = \rho[1-\langle s_u\rangle] - \eta k_u \langle s_u\rangle + \eta\sum_{v=1}^{N} A_{uv}\langle s_u s_v\rangle . \tag{3.20}$$

where $\sum_{v=1}^{N} A_{uv} = k_u$ was used. The individual-based Markov-chain approximation assumes that the joint probability of two adjacent nodes being healthy $\langle s_u s_v\rangle$ is approximately independent[5] and therefore given as

$$\langle s_u s_v\rangle \approx \langle s_u\rangle \langle s_v\rangle . \tag{3.21}$$

[5] This implies that neighboring infection state covariance vanishes, which means that they do not correlate.

Note that further assuming each node has degree $k_u \equiv n$ and that the probability of a single node being healthy is uniform for each node $\langle s_u\rangle \approx \langle s\rangle$ (as can be assumed for annealed systems) yields the mean-field result Eq. (3.16) and therefore the same epidemic threshold of $R_{0,c}$.

Instead, the implications of the actual network structure for the epidemic threshold are investigated below by assuming Eq. (3.21) in Eq. (3.20) and changing the variable of interest to the probability of being infected $\langle i_u\rangle = 1-\langle s_u\rangle$ to obtain

$$\frac{d\langle i_u\rangle}{dt} \approx -\rho\langle i_u\rangle + \eta\sum_{v=1}^{N} A_{uv}\langle i_v\rangle - \eta\sum_{v=1}^{N} A_{uv}\langle i_v\rangle\langle i_u\rangle . \tag{3.22}$$

In order to find the point of the phase transition for ρ and η, a linearization at the equilibrium $\langle i_u^\star\rangle = 0$ yields the Jacobian matrix

$$J_{uv} = -\rho\delta_{uv} + \eta A_{uv}$$

with maximum eigenvalue

$$j_{\max} = -\rho + \eta\alpha_{\max}$$

where $\alpha_{\max}$ is the largest eigenvalue of the adjacency matrix $\boldsymbol{A}$ (note that $\alpha_{\max}$ is real due to the Perron-Frobenius theorem). In order for the system to be stable in $\langle i_u^\star\rangle = 0$ and hence disease free, it is required that $j_{\max} < 0$ which implies the epidemic threshold is given by $\eta\alpha_{\max}/\rho = 1$ such that

$$R_{0,c} = \frac{\langle k\rangle}{\alpha_{\max}} \tag{3.23}$$

where R_0 refers to the mean-field basic-reproduction number $R_0 = \eta \langle k \rangle / \rho$. As stated in [52, 116], for sufficiently random networks with degree sequence k_u, the adjacency matrix's largest eigenvalue is

$$\alpha_{\max} = \frac{\langle k^2 \rangle}{\langle k \rangle} = \langle k \rangle \left(1 + \mathrm{CV}[k]^2\right)$$

where the coefficient of variation is defined as $\mathrm{CV}[k]^2 = \mathrm{Var}[k] \,/\, \langle k \rangle^2$ and quantifies the heterogeneity of the node degree. This result further arises in a different derivation where annealed contact structure is assumed with the additional condition that a node's degree follows a heterogeneous distribution [9, pp. 192–193], referred to as the heterogeneous mean field approximation. A major implication is that for random networks the epidemic threshold is given as

$$R_{0,c} = \frac{1}{1 + \mathrm{CV}[k]^2}. \tag{3.24}$$

One may interpret this as follows: In a completely regular network, where $\mathrm{CV}[k] = 0$, the epidemic threshold should correspond to the mean-field epidemic treshold $R_{0,c} = 1$. As soon as degree inhomogeneity is introduced, nodes may have both a lower and a larger degree than an average node. Eq. (3.24) implies that nodes with larger degree may pass an infection to a larger amount of other nodes and make the system more susceptible to a disease, which is not countered by the decreased susceptibility of nodes with lower degrees. In the case of BA networks in thermodynamic limit, $\mathrm{CV}[k]$ diverges and as such the epidemic threshold vanishes, making the system susceptible for any disease.

While Eq. (3.22) aims at the approximation of a Poisson process, a discrete-time approximation takes into account that multiple events may happen at once during one time step [50, 91], which increases the accuracy when numerically evaluating the individual-based Markov-chain approximation but drives the description further away from the original Poisson process, which only allows for a single event per infinitesimal increase in time. Then, the probability for a healthy node u not to be infected during a single time step is given as

$$q_u(t, \Delta t) = \prod_{v=1}^{N} \Big[1 - \eta \Delta t A_{uv} \langle i_v(t) \rangle \Big], \tag{3.25}$$

which reads as the probability of not being infected by its first neighbor *and* not being infected by the second neighbor and so forth. As such, the probability of a node to be infected after a discrete-time step Δt is

$$\langle i_u(t + \Delta t) \rangle = 1 - q_u(t, \Delta t) \times \Big[1 - \langle i_u(t) \rangle \times (1 - \rho \Delta t) \Big]. \tag{3.26}$$

Note that with $\Delta t \to dt$ and $(dt)^n \to 0$ for $n > 1$, this equation reduces to Eq. (3.22).

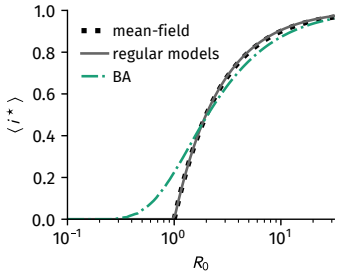


Figure 3.4: Results of the individual-based Markov-chain approximation (MCA) regarding epidemic response curves on static network models: the k-regular nearest-neighbor lattice, the random k-regular network and the Barabási-Albert (BA) model, each with a mean degree of $\langle k \rangle = 6$ and $N = 2 \times 10^4$ nodes. Shown here are equilibrium values $\langle i^\star \rangle$ obtained by integrating Eq. (3.26) for $t_{\mathrm{eq}} = 200/\rho$ with fixed recovery rate $\rho = 1$ and time step $\Delta t = 0.01/\rho$. For each value of R_0, a single network was drawn from the ensemble of the corresponding model, on which the equations were integrated. If $\sum_{u=1}^{N} \langle i_u \rangle (t) \leq 10^{-6}$, the equilibrium state was set to be equal to the disease-free state. For both regular network models, the Markov approximation yields the mean-field response curve and the epidemic threshold $R_{0,c} = 1$, as predicted by Eq. (3.24) (in regular networks, $\mathrm{CV}[k] = 0$). For the BA model, the appearance of nodes with high degree makes the system susceptible for a disease, even for lower values of R_0.

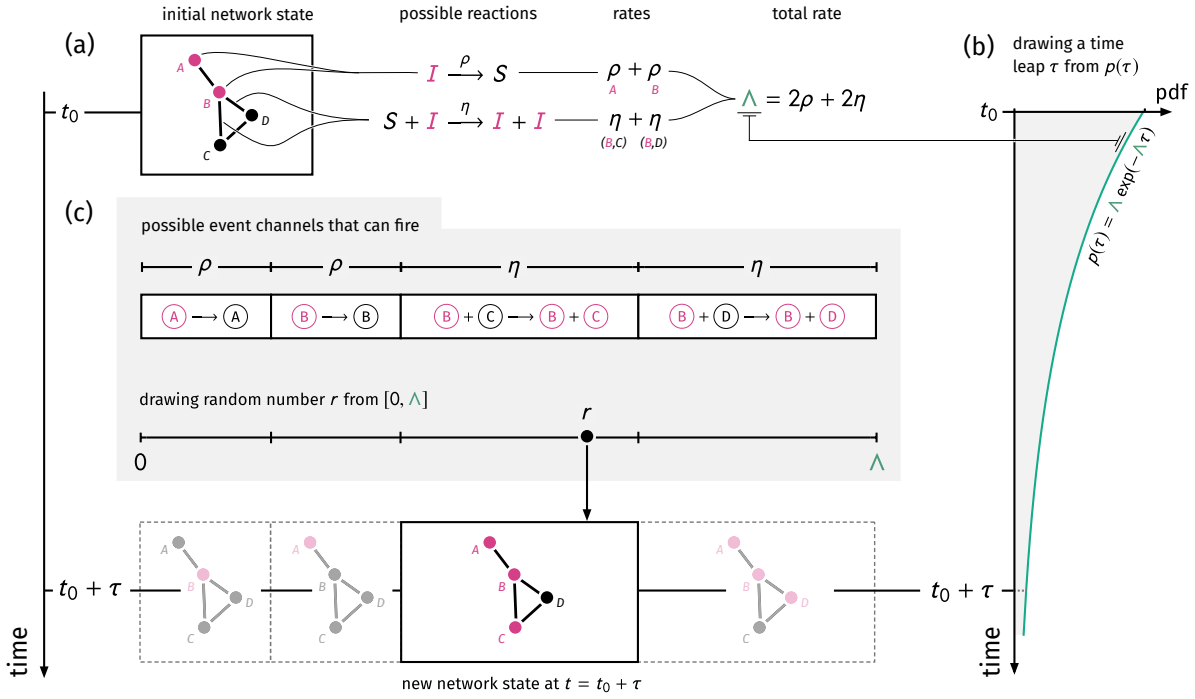


Figure 3.5: Gillespie's algorithm for SIS dynamics on static networks. **(a)** The configuration of infected and susceptible nodes on the network is described by a set of infected nodes (nodes A and B) and a set of susceptible-infected (SI) links (links (B, C) and (B, D)). Each infected node is associated with the recovery rate ρ and each SI-link is associated with the infection rate η. Hence, the total event rate is $\Lambda = 2\rho + 2\eta$. **(b)** Since the process is a Poisson process, the time-lapse τ of the first event to occur is distributed according to an exponential distribution with mean Λ^{-1}. **(c)** After drawing the time-lapse τ from this distribution, a single event is chosen from the set of all possible events. Each event occurs with probability proportional to its rate. Afterwards, the sets of infected nodes and SI-links is updated accordingly.

In Fig. 3.4, epidemic response curves obtained by integrating Eq. (3.26) are shown for three static network models: the k-regular nearest-neighbor lattice, the random k-regular network and the Barabási–Albert (BA) model, each with a mean degree of $\langle k \rangle = 6$ and $N = 2 \times 10^4$ nodes. The approximation predicts epidemic response curves equal to the mean-field result for both regular networks. For the BA model, the epidemic threshold is predicted to decrease to lower values of $R_{0,c}$—the introduction of nodes with large degrees renders the system more susceptible for a spreading disease.

3.2.3 *SIS Simulation Algorithms on Static Networks*

Here, several SIS simulation algorithms are introduced to obtain the endemic state numerically.

Gillespie's Stochastic Simulation Algorithm

The SIS process as introduced above is a Poisson process. Every infected node recovers with a constant rate ρ and every SI-link converts its corresponding susceptible constituent to an infected node with rate η. One of the conditions for a process to be Poissonian is that in an infinitesimal time step dt at maximum a single event may happen. In order to make sure that this condition is met, Gillespie introduced the so-called stochastic simulation algorithm (SSA) [73]. It is explained in detail in App. D, but briefly outlined here and illustrated in Fig. 3.5.

Considering there are a number $I(t_0)$ of infected nodes and a number $[SI](t_0)$ of links between susceptible nodes and infected nodes

at time t_0, the total rate of any event happening is given by

$$\Lambda(t_0) = \sum_j \lambda^{(j)}(t_0) = \rho I(t_0) + \eta [SI](t_0),$$

as illustrated in Fig. 3.5a. The probability that at least one of the $I(t_0) + [SI](t_0)$ possible events happens up to time $t_0 + \tau$ is given as

$$P(\tau) = 1 - \exp(-\Lambda(t_0)\tau),$$

as derived in App. D. Hence, in order to find a time τ at which the next event happens, it can simply be drawn from an exponential distribution $\tau \sim \mathcal{E}(\Lambda(t_0))$ (see Fig. 3.5b).[6] Afterwards, it has to be decided which of the events should actually happen at this time. The process is conceptually equal to drawing an event time $\tau^{(j)}$ for each possible event j with corresponding rate $\lambda^{(j)}$. Then, $\tau = \min\{\tau^{(j)}\}$ and the corresponding event is the event j associated with this minimal time-lapse τ. Going back to the original procedure, after drawing τ, the event to happen has to be drawn randomly. Each event j will be drawn with the conditional probability that its associated time-lapse was the minimal event time of all events given that it happened at time τ. This conditional probability is given as

$$\pi^{(j)} = \frac{\lambda^{(j)}}{\Lambda(t_0)}.$$

In practice, after drawing τ, a randomly chosen infected recovers with probability $\pi^{(\text{rec})} = \rho I(t_0)/\Lambda(t_0)$. Otherwise (with probability $\pi^{(\text{inf})} = 1 - \pi^{(\text{rec})}$), a randomly chosen SI-link converts its susceptible constituent to an infected node (see Fig. 3.5c). Then, the time is updated to $t = t_0 + \tau$ and the sets of infected nodes and SI-links are updated accordingly.

[6] The notation used here implies that τ follows an exponential distribution with mean $1/\Lambda(t_0)$.

Reactive Process

It is common to approximate the continuous-time contagion process using the so-called reactive process which can be computationally cheaper. In this approximation, time is discretized to pass in steps of length Δt and rates are replaced by probabilities. At each time step, in random order, every infected node tries to infect each of its susceptible neighbors with probability $\tilde{\eta} = \eta \Delta t$. Afterwards, it recovers with probability $\tilde{\rho} = \rho \Delta t$. Note that a node which recovers during a single time step may be re-infected by an infected neighbor. This process is only equal to the Poisson point process if the probabilities are sufficiently small as $\tilde{\rho} \to 0$ and $\tilde{\eta} \to 0$, because only then, the Poisson process' condition of only a single event occuring per time step is met. This was discussed more in-depth in [52].

However, studies usually fix the recovery probability to some large value [45, 49, 50, 52, 56, 57, 91] such as $\tilde{\rho} = 0.5$ or even $\tilde{\rho} = 1$ usually without providing any argumentation (e.g. in [50, 51]). While it is true that for the original Poissonian SIS process one can fix the timescale using an arbitrary rate, it is certainly not true for the reactive

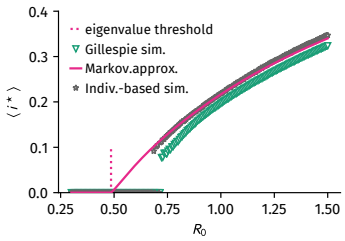


Figure 3.6: Epidemic response curves on the largest component of the author's Facebook network (introduced in Fig. 1.1)—a comparison between results obtained from exact continuous-time Gillespie simulations, the individual-based Markov-chain approximation (MCA) and the individual-based simulation (reactive process) with recovery probability $\tilde{\rho} = 1$ and recovery rate $\rho = 1$ for the Gillespie simulations (here, the time-scale does not actually matter). The MCA result was obtained by integrating Eq. (3.26) with $\Delta t = 0.01/\rho$. For the simulations, jumps in the epidemic response curves occur due to stochastic fluctuations in the finite system. The MCA and reactive process overestimate the exact response curve obtained by continuous-time Gillespie simulations. Concerning the reactive process, this is a consequence of violating the conditions of a Poisson process that only a single event per infinitesimal time step may happen. The difference between MCA and Gillespie simulation emerges for the same reason *and* because neighboring infection states are assumed to not be covariant.

process and its corresponding probabilities when they are large. As argued in [52], the reactive process will give increasingly differing results to the Poissonian process for high rates/large Δt, which is demonstrated as an example in Fig. 3.6 for an SIS process spreading on the author's Facebook friend network which was introduced in Fig. 1.1. A detailed discussion of this example is given in the unnumbered subsection titled "Comparison of the Introduced Evaluation Methods".

Quasi-Stationary State Algorithm

The SIS process has an absorbing state at $i^\star = 0$. Once all nodes are susceptible, no new infections can take place anymore. In a system of finite size ($N < \infty$) the disease may die out during SIS simulations by reaching the absorbing state randomly due to stochastic fluctuations around an endemic state $i^\star > 0$. Therefore, this non-zero endemic state $i^\star$ is referred to as the "quasi-stationary state". In order to accurately obtain the quasi-stationary state, a simulation approach was developed based on the reactive process [117]. Following this algorithm, for the first M time steps, each active configuration of infected nodes is saved (i.e. each non-zero endemic state). Then, every time a number of T time steps has passed, a configuration from the saved states is replaced by the current active states. If the current active state reaches the absorbing state, it is replaced by an active state from the saved states. It was proven in [117] that this process asymptotically converges to a set representing the ensemble of states fluctuating around the quasi-stationary state.

In this thesis, the quasi-stationary state algorithm will not be used. It was introduced here nevertheless, since the results regarding epidemic spreading on temporal networks from other studies were obtained using this algorithm.

Comparison of the Introduced Evaluation Methods

Since results regarding the epidemic response on temporal networks mostly rely on the individual-based Markov-chain approximation (MCA), its validity for accurate predictions on static networks will be tested here first. Simulation results using Gillespie's SSA and the MCA on three static network models of size $N = 2 \times 10^4$ are shown in Fig. 3.7. The MCA was evaluated by iteratively applying Eq. (3.26) with $I(0) = 100$ randomly selected initially infected nodes. The parameters used for evaluating the MCA are the same as in Fig. 3.4, i.e. $\rho = 1$ and $\Delta t = 0.01/\rho$. Tests with $\Delta t = 0.001/\rho$ did not significantly change the results of the MCA, which is why the computationally cheaper value was chosen.

As can be seen, the MCA yields satisfying results for the Barabási-Albert (BA) model and the random k-regular model of large degree. Minor differences can be seen for $\langle k \rangle = 6$ in the BA model. A larger difference arises for the random k-regular model of degree $\langle k \rangle = 6$. For this value of $\langle k \rangle$ the predicted epidemic response curve greatly

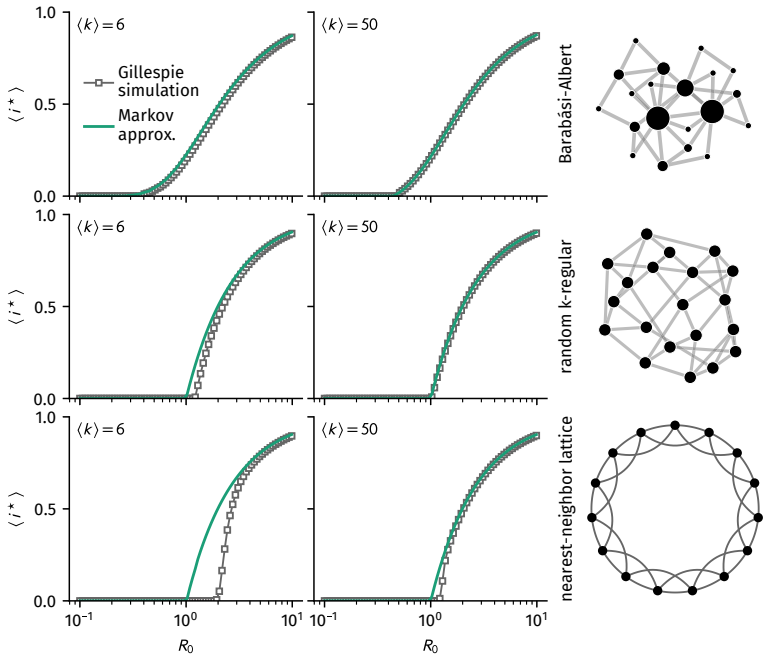


Figure 3.7: Epidemic response curves on model networks with $N = 2 \times 10^4$, obtained via Gillespie simulations and the individual-based Markov-chain approximation (MCA), iteratively integrating Eq. (3.26) with $\rho = 1$ and $\Delta t = 0.01/\rho$. Barabási-Albert (BA) and random k-regular networks are well-mixed enough for the MCA to yield similar response curves, albeit with small differences for a lower degree of $\langle k \rangle = 6$. The MCA yields larger differences for the nearest-neighbor lattices, particulary underestimating the epidemic threshold on the $k = 6$ lattice by a factor of $1/2$. These differences decrease with increasing k but are still not negligible for $k = 50$. Here, for each value of R_0 and $\langle k \rangle \in \{6, 50\}$, model networks were sampled $N_{\text{meas}} = 10$ times. On each of the samples, a disease was simulated using Gillespie's SSA as described above, beginning with $I(0) = 100$ randomly chosen initially infected nodes. The time-scale was fixed using $\rho = 1$ and the disease was considered equilibrated after $t_{\text{eq}} = 200/\rho$. Subsequently, the disease was further simulated for $t_{\text{meas}} = 100/\rho$, a sample of the number of infected taken every $\Delta t_{\text{sample}} = 1/\rho$. The endemic state was then evaluated over all simulations which did not end in the absorbing state, averaged over all samples taken during the measurement period.

differs from the simulation results. Correspondence increases for the larger value of the mean degree, however still showing slight differences for basic reproduction numbers near the epidemic threshold. As has been demonstrated in [52], infection states of neighboring nodes are highly covariant in the nearest-neighbor lattice model, rendering the assumptions of the MCA invalid for this structure. Hence, even though Δt is small, significant deviations may arise.

While there are noticeable differences between the results of the MCA and exact continuous-time simulations, studies often claim that the MCA yields nearly exact results when compared to simulations [49–51]. However, within these studies simulations are usually performed using the reactive process with large simulation probabilities per unit time step at which point the reactive process becomes essentially different from the original SIS process. An example is illustrated in Fig. 3.6 where the response curves of the MCA, the reactive process, and continuous-time Gillespie simulations are compared for SIS spreading on the largest component of the author's Facebook friend network with a time-step of length $\Delta t = 1$ and a recovery probability of $\tilde{\rho} = 1$. Since one of the time-scales can be chosen arbitrarily for continuous-time simulations, the corresponding rate was set to be $\rho = 1$ which would correspond to the MCA approximation $\tilde{\rho} \to \rho \Delta t$.[7] Varying R_0, the infection rate was computed as $\eta = R_0 \rho / \langle k \rangle$. The infection probability was computed as $\tilde{\eta} = R_0 \tilde{\rho} / \langle k \rangle$. The MCA yields satisfying predictions for the reactive process but systematically overestimates the epidemic response curve compared to the results of the continuous-time simulations, allegedly underestimating the epidemic threshold.[8] Concerning the difference between Gillespie simulations and reactive process, this is a consequence of violating the conditions of a Poisson process that

[7] Note that this approximation is actually *not* justified because for the recovery probability to be $\tilde{\rho} = 1 - \exp(-\rho \Delta t) = 1$ in time Δt considering a Poisson process, the recovery rate would have to approach $\rho \to \infty$. This further illustrates the problems arising with large probabilities in the reactive process.

[8] Note that for Gillespie and reactive process simulations, jumps in the epidemic response curves occur due to stochastic fluctuations in the finite system.

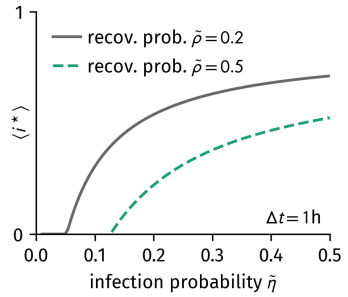


Figure 3.8: Epidemic response curves on the HT09 data set according to Valdano, *et al.* Curves were evaluated using Eq. (3.26) and Eq. (3.27) with $\eta\Delta t \to \tilde{\eta}$ and $\rho\Delta t \to \tilde{\rho}$ with $\Delta t = 1\,\mathrm{h}$. Note that the infection probability $\tilde{\eta}$ was not rescaled when $\tilde{\rho}$ changed which explains the difference between the curves. The epidemic response curves are redrawn as functions depending on the basic reproduction number in Fig. 3.9.

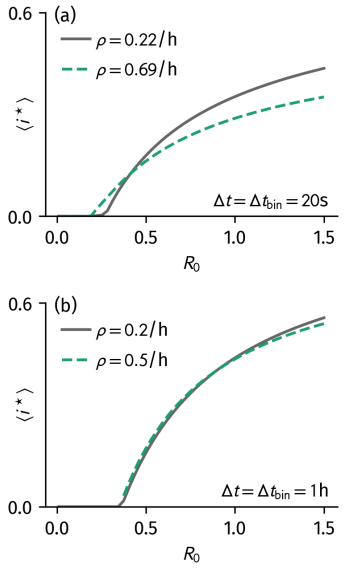


Figure 3.9: Epidemic response curves on the HT09 data set according to Valdano *et al.* [56], rescaled to be dependent on the basic reproduction number rather than the infection probability like in Fig. 3.8.

[9] Problems arise when Δt becomes too large for any large rates λ: in this case, the probability of an event happening at least once in the interval is not equal to $\lambda\Delta t$, rather, it shold be given as $\tilde{\lambda} = 1 - \exp(-\lambda\Delta t)$. In the studies cited above, large probabilities $\tilde{\eta}$ and $\tilde{\rho}$ are often used. In this case, the infection probability in Eq. (3.27) is replaced as $\eta\Delta t \to \tilde{\eta}$ and the recovery probability in Eq. (3.26) is replaced as $\rho\Delta t \to \tilde{\rho}$. Consequently, when changing Δt from the originally used value Δt_{orig} to other values one needs to transform the corresponding probabilities according to the actual rate $\lambda = -(\Delta t_{\mathrm{orig}})^{-1}\log(1-\tilde{\lambda})$ first. This is unnecessary for $\lambda\Delta t \ll 1$ at which $\tilde{\lambda} = \lambda\Delta t$ is a justified approximation.

only a single event per infinitesimal time step may happen. The difference between MCA and Gillespie simulation emerges due to the same violation *and* because neighboring infection states are assumed to not be covariant. Allegedly, the reactive process and the MCA yield similar results because multiple events happening at the same time might greatly reduce the neighboring infection state covariance, which makes the MCA applicable to model the reactive process for these parameter values.

Due to these results, caution is advised for describing epidemic spreading based on discrete-time processes with large probabilities and considering an individual-based point of view, as already argued in [52]. In Ch. 8, it will be shown that the difference between these simulation perspectives may lead to diametral results in temporal networks.

3.2.4 SIS on Temporal Networks

In recent years, a number of results regarding the prediction of the epidemic threshold and epidemic response curves on temporal networks were published, all of which were based on the individual-based Markov-chain approximation [45, 56, 57, 91]. Here, methods are introduced to obtain epidemic response curves on temporal networks and the results of the studies will be discussed.

Individual-Based Markov-Chain Approximation

Adapting the individual-based Markov-chain approximation for temporal networks, the constant adjacency matrix is simply replaced with the time-dependent one [56]. Consequently, the integrative equations Eq. (3.26) remain unchanged, however the probability $q_u(t, \Delta t)$ not to be infected by any neighboring node in the interval $[t, t+\Delta t]$ is given as

$$q_u(t,\Delta t) = \prod_{v=1}^{N} \left[1 - \eta\Delta t A_{uv}(t)\,\langle i_v(t)\rangle\right]. \tag{3.27}$$

Here, Δt is the constant length of a time step, usually the duration of a discrete-time temporal network's single snapshot duration Δt_{bin} or any value $\Delta t = \Delta t_{\mathrm{bin}}/n$ which devides this snapshot duration into n equally-sized intervals. For continuous-time temporal networks, this approach is unfeasible because there exists no value Δt which would evenly divide *every* snapshot duration, as those follow a continuous distribution function.

Therefore, in this work, an adapted continuous-time integration approach will be chosen for continuous-time temporal networks whose adjacency matrices may have such an aribitrary duration. Within this approach, a small maximum MCA integration time of $\Delta t_{\max}$ is introduced to ensure that the approximation $\tilde{\lambda} \approx \lambda\Delta t$ is always justified for any rate λ.[9] Suppose a system is currently at time t and the next change of the adjacency matrix will happen at time $t_{\mathrm{next}} > t$. Then the integration time-step is chosen as $\Delta t = \min(\Delta t_{\max}, t_{\mathrm{next}} - t)$.

Consequently, Eq. (3.27) and Eq. (3.26) are evaluated with this value of Δt and time is advanced as $t \to t + \Delta t$. This approach is equal to the discrete-time approach if the maximum integration step $\Delta t_{\max}$ is set to the length of a single discrete-time snapshot duration Δt_{bin}.

Using the reactive process and the individual-based Markov-chain approximation (MCA), Valdano *et al.* showed that the MCA accurately predicts reactive process simulation results for a collection of real-world temporal networks [56]. They further developed an analytical method to compute the epidemic threshold on temporal networks, based on the maximum eigenvalue of the so-called infection propagator which is evaluated from adjacency matrices based on Eq. (3.26) and Eq. (3.27). This method, however, will not be of further concern, since interest lies on the epidemic response curve here. Evaluating the full response curve comes with the advantage of a concurrent estimation of the epidemic threshold.

The HT09 data set was one of the temporal networks for which the epidemic response curve was evaluated in [56]. Instead of using the high resolution data with $\Delta t = 20\,\text{s}$, the authors discarded nightly data and the remaining data was rebinned to a time of $\Delta t_{\text{bin}} = 1\,\text{h}$. An epidemic was then simulated with discrete time of step length $\Delta t = \Delta t_{\text{bin}}$, choosing recovery probabilities per time step $\tilde{\rho} \in \{0.2, 0.5\}$ and infection probabilities per time step $\tilde{\eta} \in [0, 0.5]$. The full epidemic response curves were obtained using the MCA and the quasi-stationary state algorithm. In the temporal variant of the quasi-stationary state algorithm, for each snapshot, a number of active configurations was saved to be used as replacements when the disease dies out in this particular snapshot. The results of the MCA (which were found to be highly accurate with respect to the simulations) are shown in Fig. 3.8, which is a reproduction of Fig. 2d in [56]. At first sight one may conclude that increasing the recovery probability $\tilde{\rho}$ increases the epidemic threshold and lowers the number of infected, this effect, however, is simply explained by the fact that $\tilde{\eta}$ was not rescaled when varying $\tilde{\rho}$. Even in the mean-field case, this would yield a change in the basic reproduction number and therefore a change in the epidemic response and epidemic threshold. The curves should therefore be compared as functions depending on the basic reproduction number $R_0 = \eta \langle k \rangle / \rho$, as shown in Fig. 3.9b. As one may observe in this figure, the responses are actually not substantially different but there is a slight decrease in the epidemic threshold. This deviation is even more pronounced if the network is not rebinned but left in its original state with resolution $\Delta t = 20\,\text{s}$. As shown in Fig. 3.9a, the epidemic response curves are different on the temporal network of higher temporal resolution, essentially revealing that changing the temporal resolution for both integration and structural variation changes how a disease spreads in the system. With the smaller temporal resolution, it is shown that the epidemic threshold decreases with increasing recovery rate. For both rates, the epidemic threshold decreases compared to the rebinned case with $\Delta t = 1\,\text{h}$.[10]

[10] Note that as discussed above, the recovery probabilities are recomputed for $\Delta t = 20\,\text{s}$ to obtain rates reflecting the corresponding probabilities.

Speidel *et al.* [45] argued that coarse-graining temporal change in the SIS dynamics to long times such as $\Delta t = 1\,\mathrm{h}$ might change the results. They further noted that while the theoretical approach by Valdano *et al.* yields satisfying predictions, it is still unclear in which way the properties of temporal networks influence the epidemic response curves for varying recovery rates. Therefore, they introduced what they call a "continuous-time" approach to study the change of the epidemic threshold for varying recovery rates. In this approach, they replace probabilities with rates and linearize the Markov-chain approximation equation at the disease-free state $i^\star = 0$.[11] Consequently, they find the epidemic threshold as the maximum eigenvalue of an infection-propagator similar to the one defined in [56]. Using this approach and quasi-stationary state simulations, they show that in all of the empirical temporal networks as well as almost all of their investigated model temporal networks, the epidemic threshold decreases when the recovery rate increases for fixed R_0 (see Fig. 3.10 for an example result on the HT09 data set). In the limit $\rho \to 0$, a recovered node 'sees' the whole temporal network several times before recovering, hence the epidemic response behaves as if the process was run on the averaged static network. For larger recovery rates, an infected node experiences a limited number of structural changes and hence the temporal structure becomes more important. Based on their results they claim that networks become more susceptible to diseases the more "temporal" they are. They find a single exception for a model which consists of isolated nodes and a number of isolated pairs in each snapshot (cf. Sec. 2.4.3). In this case, the epidemic threshold *increases* with increasing recovery rate making the system *less* susceptible for diseases, while the MCA predicts a decrease. The reason for this discrepancy is that in this highly sparse structure, neighboring infection states are highly covariant and therefore the individual-based approximation breaks down (for an explicit example, see App. E.5).

[11] Note that recently, a similar "continuous-time" approach was taken, where in the individual-based Markov-chain approximation, the limit $\Delta t \to dt$ yields a system of ODEs [57].

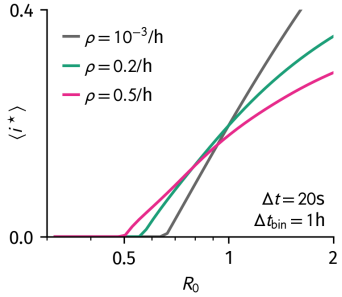


Figure 3.10: Epidemic response curves on the HT09 data set according to Speidel *et al.* [45], evaluated based on the individual-based Markov chain approximation obtained by integrating Eq. (3.26) and Eq. (3.27) with $\Delta t = 20\,\mathrm{s}$. Note that again, the analyzed data set was binned to an aggregation window of $\Delta t_{\mathrm{bin}} = 1\,\mathrm{h}$ without further reasoning. The epidemic threshold decreases with increasing recovery rate.

Unfortunately, it is unclear in which way their simulations are supposed to be of continuous time. The authors state that they use a quasi-stationary state algorithm similar to the one used by Valdano *et al.*, supposedly with an integration step much smaller than the one used by Valdano *et al.*, but not further specified. They state that they divide each snapshot of length Δt_{bin} in two halves and assign active states to each of the halves in case the process reaches the absorbing state in either one of those halves. They do not state how they computed transition probabilities for a single step, however they refer to the original quasi-stationary state algorithm which computes probabilities as $\tilde{\lambda} = \lambda \Delta t$. If in fact, the reactive process was used for simulation, the approach is not of continuous-time if Δt was chosen too large, in which case the reactive process differs from the exact Poisson process. Furthermore, for reasons not further specified, they change the structure of all investigated temporal networks by rebinning to larger aggregation windows (for instance $\Delta t_{\mathrm{bin}} = 1\,\mathrm{h}$ for HT09). In its original form, the HT09 is structurally

very close to their investigated model of isolated pairs and isolated nodes in which the Markov-chain approximation is invalid. Rebinning substantially changes the single snapshot structure to a point at which the investigated substrate is completely different from the original network.

For these reasons, it is essential to reevaluate and compare epidemic observables on the original, highly resolved temporal networks as well as on continuous-time temporal network models. These evaluations should be done with an exact continuous-time simulation algorithm (introduced below) and compared to evaluations using the Markov-chain approximation, both of which will be done in Ch. 8.

Gillespie's Algorithm for Temporal Networks

In order to simulate the exact continuous-time SIS process, a Gillespie algorithm for temporal networks will be used, first introduced in [89]. At the basis of this algorithm lies the formulation of an inhomogeneous Poisson process with time-dependent rate $\Lambda(t) = \sum_j \lambda^{(j)}(t) = \rho I(t) + \eta[SI](t)$. In the context of temporal networks, $\Lambda(t)$ may change as a step function every time the adjacency matrix changes and consequently the number of SI-links. For such an inhomogeneous Poisson process, given that the last event happened at time t_0, the probability of at least one event happening up to time τ is

$$P(\tau) = 1 - \exp\left(-\int_{t_0}^{t_0+\tau} dt\, \Lambda(t)\right).$$

Consequently, the dimensionless variable

$$\theta(\tau) = \int_{t_0}^{t_0+\tau} dt\, \Lambda(t) \tag{3.28}$$

is exponentially distributed with unit mean as $\theta \sim \mathcal{E}(1)$. Hence, the time τ until the next event can be found by randomly drawing θ from this exponential distribution and then numerically solving Eq. (3.28) for τ. In the case of $\Lambda(t)$ changing as a step function, $\theta(\tau)$ is a collection of piece-wise defined linear functions and therefore can be solved exactly. As soon as the time τ is found, the event channel j triggering this event is chosen at random with probability

$$\pi^{(j)} = \frac{\lambda^{(j)}(t_0+\tau)}{\Lambda(t_0+\tau)}.$$

The exact workings of the algorithm as well as the derivation of these formulas is shown in App. D.2 and an exhaustive explanation is given in [89].

Part II

Random Walks on Modular Hierarchical and Other Small-World Networks

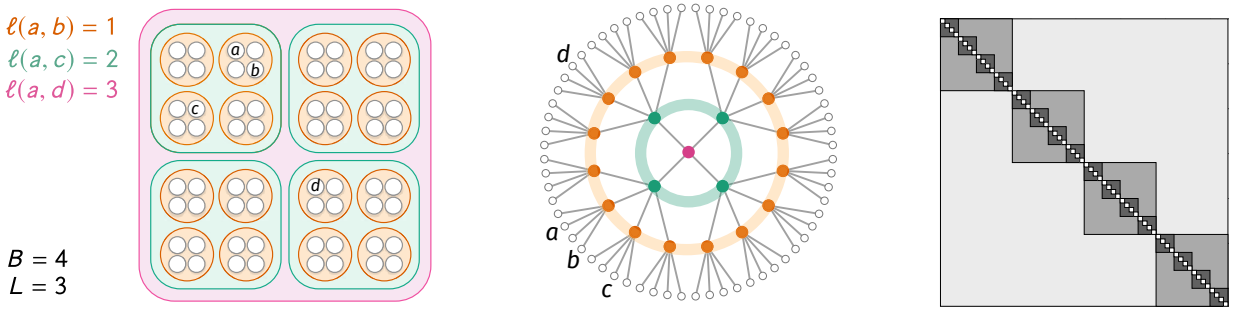
4
Small Worlds Are All the Same

The aim of the second part of this thesis is to test the hypothesis whether modular hierarchical networks are in a structural optimum for random walk processes. One way to pursue this is to devise a synthetic model whose properties are closely related to the modular hierarchical structures observed in empirical networks. This structure may then be altered using a single control parameter to allow for comparisons to different known structures. Such a model, as we introduce in this chapter, is closely related to the modular hierarchical Watts-Newman network model [13] summarized in Sec. 2.2.4. Since this original model was developed to explain the small-world phenomenon, the discussion will be extended to a variety of small-world network models which all follow the same principles listed below.

1. By adjusting a single control parameter, one interpolates between generative network models which produce locally clustered structures with high regularity and the Erdős–Rényi network model.

2. During interpolation the mean degree remains constant. Since the mean degree is a first-order control parameter for certain dynamic processes, this allows for a thorough comparison of the structural consequences beyond the bias introduced by a varying total number of edges in the network.

While each of the models devised in the following has been introduced in other variants before (see Sec. 2.2.4), the defining feature here will be that while conceptually different, their new definitions will allow for a thorough map between their control parameters which will reveal how, despite their apparent differences, their structural properties are very similar, finally developing a generalized model. It will further be shown that the conceptual differences of the models do not yield too different results when the flow-based modular hierarchical partition algorithm *Infomap* is applied, which suggests that all of these small-world models produce structures which can be interpreted as being modular hierarchical. Large parts of this chapter are very similar to texts already published by the author in [93, 94]. These parts will be explicitly referenced.



(a) (b) (c)

Figure 4.1: Hierarchical modular structure as introduced in Sec. 4.1, using a number of modules per level (base) $B = 4$ and total number of levels $L = 3$ as an example. **(a)** and **(b)** represent the hierarchy of modules as embedded structures and a hierarchical tree, respectively. The network nodes (white circles) are grouped into modules of B components at level 1 (orange), which are in turn grouped into B-sized modules (containing B^2 nodes) at level 2 (turquoise). Finally, these are grouped into a single level 3 structure (pink) that also contains B submodules, and thus B^3 nodes in total. The nodes labeled a and b belong to the same level-1 module, those labeled a and c are in the same level-2 module (but not the same level-1 module), and those labeled a and d, in the same level-3 module (but not the same lower-level ones). A hierarchical modular network can then be built by linking with lower probability the nodes that belong to higher level modules only. **(c)** The structure of the resulting adjacency matrix, with darker shades of grey representing a higher probability of having '1' entries that denote connections. The figure was reproduced from [93].

4.1 *Self-Similar Modular Hierarchical Random Networks*

In this section a stochastic block model is presented that can generate self-similar random networks with varying 'intensity' of modular hierarchical structure (where 'intensity' refers to the amount of short-ranged connections an average node has, as will be indicated). It is a variant of the model introduced by Watts *et al.* [13] briefly discussed in Sec. 2.2.4. Its definition will be shown and some of its network properties will be derived. This section is close to the text first published in [93].

4.1.1 *Model Definition and Structural Phases*

The model was developed in collaboration with C. Huepe and D. Brockmann. Similar to the model by Watts, Dodds, and Newman [13], it assumes that each node belongs to a module that is part of a bigger community of modules, which is in turn part of a yet larger community of modules of modules, etc., as depicted in Fig. 4.1a. This structure of communities and subcommunities can be represented by an underlying hierarchical tree (see Fig. 4.1b), where the *hierarchical distance* ℓ between two nodes is defined as the smallest number of levels needed to traverse up the tree to find a common ancestor. Note that only leaves of the tree represent actual nodes in the generated network while the rest of the tree is only used to define a modular hierarchical connectivity structure. The probability of having a connection between two nodes decreases as their distance ℓ increases.

For simplicity, and in contrast to the original model presented in Sec. 2.2.4, pure self-similar modular hierarchical (SSMH) structures are considered, where all modules have the same number of submodules. Thus, SSMH networks are generated starting from a B-ary tree of height L, where B is the module size and L is the total number of hierarchical layers. Then the final network consists of $N = B^L$ nodes. Each of those could potentially be connected to up to

$$k_\ell^{\max} = B^\ell - B^{\ell-1} = B^{\ell-1}(B-1) \tag{4.1}$$

other nodes at hierarchical distance ℓ. Although this number of potential links to other nodes grows exponentially with ℓ, in order to generate MH structures it seems intuitive to impose a limiting condition in a way that the number of connections to nodes in other modules is typically smaller than, but of similar order as, the number of connections to nodes within the same module. To this end, a connection probability P_ℓ is chosen that decreases exponentially with increasing hierarchical distance, given by

$$P_\ell \propto \left(\frac{\xi}{B}\right)^{\ell-1}. \tag{4.2}$$

Here, $0 \leq \xi \leq B$ is defined as the modular hierarchical *structural control parameter*. Hence, choosing different values of ξ allows for the generation of classes of networks with different degrees of hierarchical modularity. As argued before, the mean degree should remain constant for all values of ξ. To ensure this, one first has to find the mean degree of each node with respect to all other nodes at hierarchical distance ℓ, which is given by $\langle k_\ell \rangle = P_\ell k_\ell^{\max}$. The total mean degree is then $\langle k \rangle = \sum_{\ell=1}^{L} \langle k_\ell \rangle$, which can subsequently be used to normalize the total number of connections, obtaining

$$P_\ell = \frac{\langle k \rangle}{B-1}\left(\frac{1-\xi}{1-\xi^L}\right)\left(\frac{\xi}{B}\right)^{\ell-1}. \tag{4.3}$$

Note that if $\langle k \rangle > B-1$, the modular hierarchical structural control parameter ξ can only be chosen to be larger than or equal to $\xi_{\min} > 0$, where $(B-1)\left(1-\xi_{\min}^L\right) = \langle k \rangle (1-\xi_{\min})$, in order to have a connection probability that satisfies $P_\ell \leq 1$ for all hierarchical distances $\ell \geq 1$.

In the case of $\xi = 0$, the network consists of B^{L-1} densely connected Erdős–Rényi random networks of B nodes and connection probability $p = \langle k \rangle / (B-1)$ (note that this is, naturally, only possible if $\xi_{\min} = 0$, as indicated in the last paragraph). As ξ is increased, links are redistributed from the lowest hierarchical layer to higher ones, while keeping the total number of links constant. The system thus goes through the following structural phases:

- $\xi < \xi_c$: The network does not yet have a largest component of size $\approx B^L$. A numerical analysis reveals that for $\langle k \rangle \geq 4$ and $L \geq 3$, this critical parameter is given[1] as $\xi_c \approx B^{-1}$. Thus, it is always smaller than ξ_{SW} as defined for the next phase. An exact determination of ξ_c is left for future research.

- $\xi_c < \xi < \xi_{SW}$: The network is in a phase with *strong* hierarchical modularity, in the sense that it (i) has a largest component of size $\approx B^L$ *and* (ii) an average node has more links to nodes in its own lowest hierarchy module (level $\ell = 1$) than to nodes in all other hierarchy layer groups $\ell > 1$ combined, satisfying

$$\langle k_1 \rangle > \langle k_2 \rangle + \cdots + \langle k_L \rangle. \tag{4.4}$$

[1] This can be motivated as follows. In the highest hierarchy layer L, there are B submodules. Each submodule contains B^{L-1} nodes and thus can potentially have $B^{L-1} \times (B^L - B^{L-1})$ connections to other submodules. With $\xi = B^{-1}$, the total number of connections to other submodules is $k_{\text{sub}} = P_L(\xi = B^{-1}) \times B^{L-1} \times (B^L - B^{L-1})$ (using Eq. (4.3)). Hence $k_{\text{sub}} = \langle k \rangle \times (1 - B^{-1})/(1 - B^{-L}) \approx \langle k \rangle$. This means that on average $\langle k \rangle$ edges are leaving out of a single submodule to connect to other submodules in the highest layer L. Consequently, the larger the value of $\langle k \rangle$, the larger the likelihood that all B submodules are connected to build a large component of size $\approx B^L$ for this value of ξ.

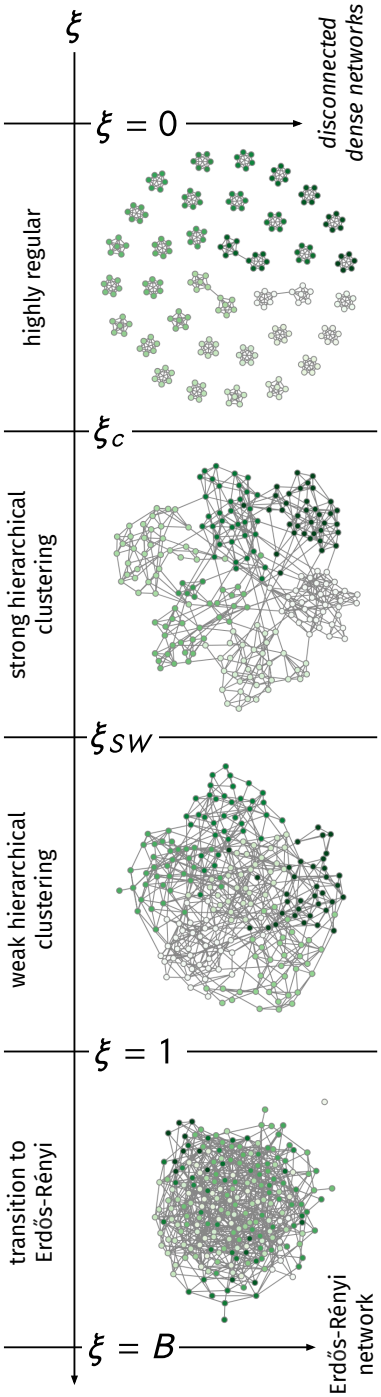


Figure 4.2: SSMH example networks for the modular hierarchical phasesdefined via Eqs. (4.4–4.6). Used parameters were $B = 6$, $L = 3$, and $\langle k \rangle = 6$.

The value of ξ_{SW} is given by the solution to equation $0 = \xi_{SW}^L - 2\xi_{SW} + 1$ (with $1/2 \leq \xi_{SW} \leq 1$), which quickly approaches $\xi_{SW} = 1/2$ as L is increased.

- $\xi_{SW} < \xi < 1$: The network is in a phase with *weak* hierarchical modularity, in the sense that the degree at each hierarchical level is smaller than the degree at the next one, that is

$$\langle k_1 \rangle > \langle k_2 \rangle > \cdots > \langle k_L \rangle \tag{4.5}$$

  is satisfied, but the condition in Eq. (4.4) is not.

- $1 < \xi < B$: In this phase, only the basic hierarchical modularity probability condition

$$P_1 > P_2 > \ldots > P_L \tag{4.6}$$

  is satisfied, as the network transitions to a homogeneous random structure.

- $\xi = B$: The network is identical to an Erdős–Rényi random network, as P_ℓ is equal for all layers ℓ.

Example networks for these different phases are displayed in Fig. 4.2.

In order to efficiently generate SSMH networks, an $\mathcal{O}(m)$ algorithm (where m is the total number of edges) is briefly outlined as follows. One begins with the construction of B^{L-1} Erdős–Rényi networks, each with number of nodes B and with connection probability P_1, using the algorithm described in [74]. Those will be the base modules in layer 1. Subsequently, for every layer $\ell > 1$, one draws a number m_ℓ of edges in this layer from a binomial distribution $m_\ell \sim \mathcal{B}(m_\ell^{\max}, P_\ell)$ with parameters $m_\ell^{\max} = B^L B^{\ell-1}(B-1)/2$ and success probability P_ℓ. For every edge appearing in the layer, a random node u is picked from the set of all nodes, followed by picking a second node v from all $B-1$ modules that this node u can reach in this layer. If there is not yet an edge connecting u and v, the edge is assigned, otherwise a new originating node u is picked. Implementations of this algorithm are publicly available for download [118].

4.1.2 *Network Properties*

In this subsection, the number of two-stars per node, number of triangles per node, degree variance and clustering coefficient of this network model are given. Especially the latter two will become of importance to explain an emergent minimum in the pair-averaged first passage time in Ch. 5.

Degree Variance

The second moment of the degree distribution is given by

$$\left\langle k^2 \right\rangle = \left\langle \left(\sum_{\ell=1}^{L} k_\ell \right)^2 \right\rangle$$
$$= \sum_{\ell=1}^{L} \left\langle k_\ell^2 \right\rangle + 2 \sum_{\ell=2}^{L} \sum_{m=1}^{\ell-1} \underbrace{\langle k_\ell k_m \rangle}_{=\langle k_\ell \rangle \langle k_m \rangle} .$$

Since the layer degrees are binomially distributed, the moments of each hierarchical layer are

$$\langle k_\ell \rangle = B^{\ell-1}(B-1)P_\ell$$
$$\left\langle k_\ell^2 \right\rangle = B^{\ell-1}(B-1)P_\ell \left(1 + B^{\ell-1}(B-1)P_\ell - P_\ell\right).$$

With Eq. (4.3), one obtains

$$\left\langle k^2 \right\rangle = \langle k \rangle + \langle k \rangle^2 \left[1 - \frac{1}{B-1} \left(\frac{1-\xi}{1-\xi^L} \right)^2 \left(\frac{1-(\xi^2/B)^L}{1-\xi^2/B} \right) \right], \quad (4.7)$$

from where the variance of the degree is found to be

$$\mathrm{Var}[k] = \langle k \rangle - \frac{\langle k \rangle^2}{B-1} \left(\frac{1-\xi}{1-\xi^L} \right)^2 \left(\frac{1-(\xi^2/B)^L}{1-\xi^2/B} \right), \quad (4.8)$$

which is a monotonically increasing function of the structural control parameter and displayed in Fig. 4.3.

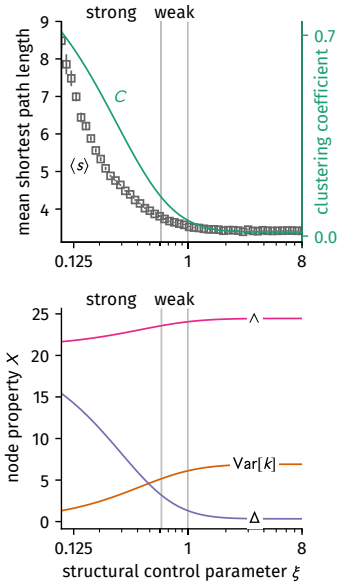


Figure 4.3: Properties of the SSMH network model for increasing control parameter ξ, here with parameters $B = 8$, $L = 3$, and $\langle k \rangle = 7$. **(Top)** The mean shortest path length $\langle s \rangle$ (as an average over the largest components from 10 independently sampled networks for each value of ξ) decreases faster than the clustering coefficient C (given via Eq. (4.11) with increasing structural control parameter ξ). **(Bottom)** The mean number of triangles per node Eq. (4.9), the mean number of two-stars per node Eq. (4.10), and the node degree variance Eq. (4.8). Note that none of the observables change significantly after leaving the weak modular hierarchically clustered phase at $\xi = 1$.

Clustering Coefficient

In the SSMH model, each node is statistically equivalent. Hence, in order to calculate the clustering coefficient as given by its definition in Sec. 2.1.4, where $C = \triangle/\wedge$, one may compute the mean number of triangles $\triangle$ per node and the mean number of two-stars $\wedge$ per node without loss of generality for a single focal node which is set to be $i = 1$. The number of triangles per node is then given as

$$\triangle = \sum_{j>1}^{N-1} \sum_{u>j}^{N} P_{\ell(1,j)} P_{\ell(1,u)} P_{\ell(j,u)}.$$

Hence, for every pair of nodes (j, u) with $j \neq u \neq 1$, one calculates the probability that node 1 is connected to node j, node 1 is connected to node u, and node j is connected to node u. Analogously, the expected number of two-stars is the sum over all pairs (j, u) of the probability that node 1 is connected to j and 1 is connected to u (thus building a two-star),

$$\wedge = \sum_{j>1}^{N-1} \sum_{u>j}^{N} P_{\ell(1,j)} P_{\ell(1,u)}.$$

Instead of summing over all node pairs, summing over pairs of hierarchical layers is preferred, as this significantly reduces the complexity of the calculation. Thus, both sums can be split into two contributions each, as

$$\triangle = \pi_S^{(3)} + \pi_L^{(3)} \quad (4.9)$$
$$\wedge = \pi_S^{(2)} + \pi_L^{(2)}, \quad (4.10)$$

where subscripts represent (*S*)hort-range and (*L*)ong-range contributions, which are all derived analytically in App. B.1. As the clustering coefficient is given as $C = \triangle/\wedge$, using Eqs. (4.9) and (4.10) yields

$$C = \frac{\pi_S^{(3)} + \pi_L^{(3)}}{\pi_S^{(2)} + \pi_L^{(2)}}. \tag{4.11}$$

It is a decreasing function of the structural control parameter ξ and shown in Fig. 4.3, alongside the mean number of triangles per node $\triangle$ and the mean number of two-stars per node $\wedge$.

4.2 *Modified Watts–Strogatz Small-World Networks*

Since a modular hierarchical network model has been introduced first to explain the small-world effect in social networks [13], it seems appropriate to also study the original small-world model which only makes distinctions between short-range and long-range connections in order to study the structural and dynamical differences of the models. This section redefines a variant of the Watts–Strogatz model which happens to be structurally equivalent to a modified model already introduced by Song and Wang [119]. It is, however, redefined below in a slightly differing manner, to suit the needs of comparison to the other models introduced in this chapter. This section is close to the text published in [94].

4.2.1 *Structural Definition and Sampling Algorithm*

As described in Sec. 2.2.4, in the original model N nodes are positioned equidistantly on a ring and subsequently *locally* connected, i.e. connected to nodes in their vicinity with maximal lattice distance $d \leq k/2$ where k is an even positive integer and equal to the degree of every node. For the rewiring process, each node rewires its connections to its $k/2$ rightmost neighbors to any other node in the network with probability p_r.

Within the adapted model defined below, edges posses an inherent probability to exist, which varies for *short-range* (*S*) and *long-range* (*L*) contacts, which already draws similarities to the SSMH model devised above. A potential contact between nodes (i, j) is considered to be short-ranged if their distance in periodic boundary conditions is $d(i, j) \leq k/2$; it exists with probability p_S. It is considered long-range if $d(i, j) > k/2$ and exists with probability p_L. The distance is computed as $d(i, j) = \min(|j - i|, N - |j - i|)$. In short, two nodes with lattice distance d are connected with probability

$$p_d = \begin{cases} p_{S,} & \text{if } d \leq k/2, \\ p_L, & \text{otherwise.} \end{cases}$$

Hence, if $p_S = 1$ and $p_L = 0$, the model produces a structure which is equal to the original model's starting point, a one-dimensional k-nearest neighbor lattice. On the other hand, if $p_S = p_L \equiv p$, each

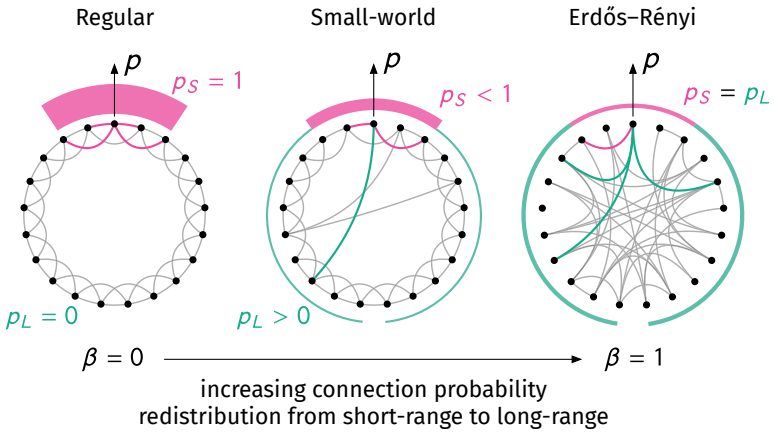


Figure 4.4: Schematic representation of the modified small-world model as defined in this section. Much like in the original model, the underlying structure is composed of N nodes placed equidistantly on a ring. However, instead of rewiring a previously defined lattice, each pair of nodes is connected with distance-based probability p_d where d is their minimal distance on the ring. Within distance $d \leq k/2$, nodes are connected with short-range probability p_S. For larger distances, nodes are connected with long-range probability $p_L = \beta p_S$. With increasing redistribution parameter $0 \leq \beta \leq 1$ connection probability is redistributed from the short-range regime to the long-range regime while the mean degree k is kept constant. Hence at $\beta = 0$ the short-range probability is unity while the long-range probability is zero which produces a k-nearest neighbor lattice. With increasing β, long-range "short-cuts" become more probable until at $\beta = 1$ both connection probabilities are equal and thus the model becomes equal to the Erdős–Rényi model. The figure was taken from [94].

edge exists with probability p and hence the model is equal to the Erdős–Rényi network model. The mean degree is held constant by noticing that it is composed of a short-range degree $\langle k_S \rangle$ and a long-range degree $\langle k_L \rangle$. Each node has k potential short-range neighbors and $N - 1 - k$ potential long-range neighbors. Thus, its expected degree is $p_S k + p_L(N - 1 - k) = \langle k_S \rangle + \langle k_L \rangle = \langle k \rangle \equiv k$. To keep the mean degree constant, a *structural control parameter* β is introduced which controls the trade-off of connection probability in the short- and long-range regimes such that $p_L = \beta p_S$. Note that at $\beta = 0$, we have $p_L = 0$ and $p_S = 1$ while at $\beta = 1$ we find $p_L = p_S \equiv p$. In order for the mean degree to be constant, the distance-based probabilities are evaluated to be

$$p_S(\beta) = \frac{1}{1 + \beta(N - 1 - k)/k}, \tag{4.12a}$$

$$p_L(\beta) = \frac{\beta}{1 + \beta(N - 1 - k)/k} = \beta p_S(\beta). \tag{4.12b}$$

The short-range node degree k_S follows a binomial distribution $\mathcal{B}(k, p_S)$ and the long-range node degree k_L follows a binomial distribution $\mathcal{B}(N - 1 - k, p_L)$. A schematic explanation of the model is given in Fig. 4.4. A simple network generation algorithm is given as follows. Each node $0 \leq u \leq N - 1$ connects to each of its $k/2$ rightmost short-range neighbors with probability p_S. Afterwards, m_L long-range edges are drawn, where m_L follows $\mathcal{B}(N(N - 1 - k)/2, p_L)$. For each long-range edge, one chooses a source node u uniform at random from $[0, N - 1]$. This node is then connected to a long-range neighbor $v = (u + k/2 + z) \mod N$ where the integer z is drawn uniform at random from the interval $[1, N - k - 1]$. If an already existing edge was chosen, repeat the procedure for this long-range edge. This algorithm has complexity $\mathcal{O}(Nk + \langle m_L \rangle)$ for sparse networks. Implementations of the algorithm are available as open source C++/Python packages [118, 120].

Example networks for increasing structural control parameter are displayed in Fig. 4.5.

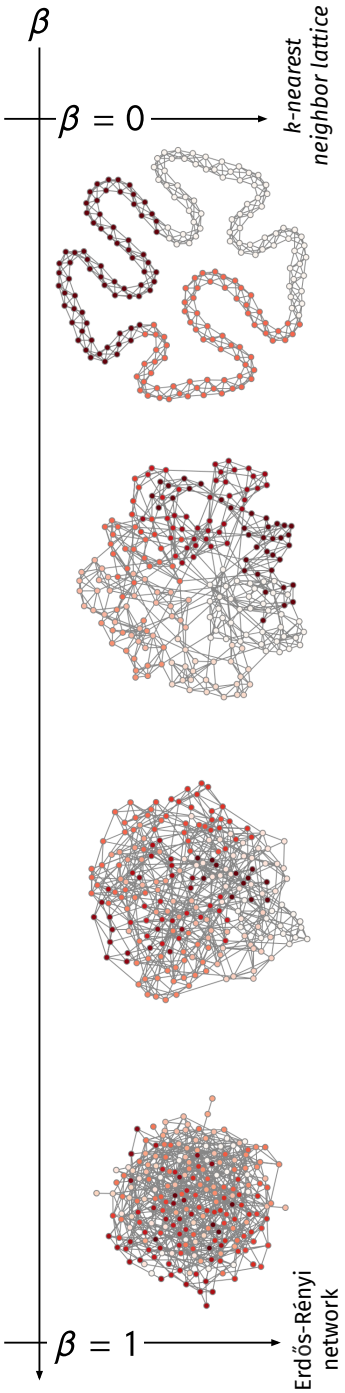


Figure 4.5: Network samples from the modified Watts–Strogatz model for increasing structural control parameter β. Parameters $N = 6^3$ and $k = 6$ were used.

4.2.2 *Network properties*

Since the node degree is given as the superposition of short-range and long-range degree, the degree variance can be found as

$$\mathrm{Var}\left[k\right] = \mathrm{Var}\left[k_S\right] + \mathrm{Var}\left[k_L\right] = kp_S(1-p_S) + (N-1-k)p_L(1-p_L). \tag{4.13}$$

For increasing β both short-range and long-range variances increase, such that the degree variance is an increasing function of β, as shown in Fig 4.7b. The full degree distribution is evaluated by noting that any node degree is $k_i = k_{S,i} + k_{L,i}$, such that its distribution evaluates to

$$\begin{aligned} p_{k'} &= \sum_{k_L=0}^{\infty}\sum_{k_S=0}^{\infty} P_S(k_S)P_L(k_L)\delta_{k',k_S+k_L} \\ &= \sum_{k_S=0}^{\min(k',k)} \binom{k}{k_S}\binom{N-1-k}{k'-k_S}(1-p_S)^{k-k_S}\times \\ &\qquad \times p_S^{k_S}(1-p_L)^{N-1-k-k'+k_S}p_L^{k'-k_S}, \end{aligned} \tag{4.14}$$

which is virtually equal to the result derived in [119] (the only difference being that they let the sum run to k) and is shown in Fig. 4.7a.

The following evaluation of the clustering coefficient will rely on its definition in Sec. 2.1.4 where $C = \triangle/\wedge$. To evaluate the expected number of two-stars per node $\wedge = \langle A_{iu}A_{iv}\rangle$ one observes that a node of degree k_v is part of $(1/2)k_v(k_v-1)$ two-stars such that the mean number of two-stars per node is given as

$$\wedge = (1/2)\left[\mathrm{Var}\left[k\right] + k(k-1)\right], \tag{4.15}$$

which is illustrated in Fig. 4.7b.

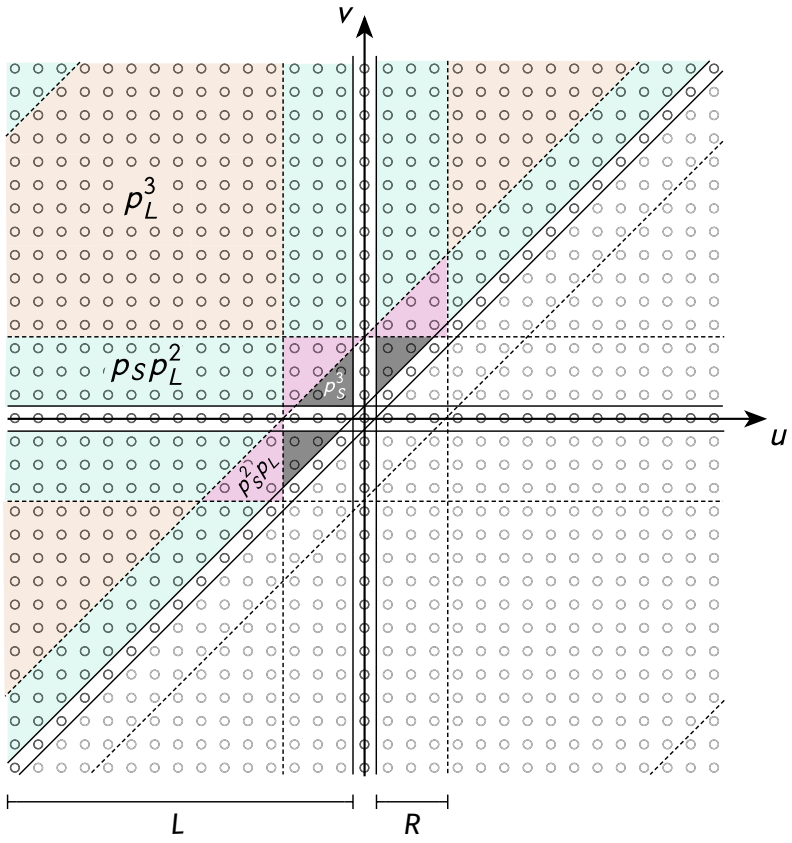


Figure 4.6: Schematic figure for the evaluation of the areas of summation to find the expected number of triangles per node $\triangle$ for odd numbers of nodes N as per Eq. (4.16). Note that here, the sum has been shifted to be $\sum_{u=-(N-1)/2+1}^{(N-1)/2} \sum_{v=u+1}^{(N-1)/2}(\cdot)$ such that u and v are equal to their lattice distance to a focal node at $d = 0$. Straight lines mark the border of the areas within which $u \neq v \neq 0$, $|u| \leq k/2$, and $|v| \leq k/2$, and dashed lines mark the border of areas where $|d(u,v)| \leq k/2$. The figure was taken from [94].

In order to find the expected number of triangles per node $\triangle$ one recognizes again that every node is statistically equivalent. Thus, without loss of generality, one may compute the number of triangles per node for a single focal node $i = 1$ based on the definition in Sec. 2.1.4 as

$$\begin{aligned} \triangle &= \sum_{u=2}^{N-1} \sum_{v=u+1}^{N} p_{d(u,1)} p_{d(v,1)} p_{d(u,v)} \\ &= F p_S^3 + G p_S^2 p_L + H p_S p_L^2 + I p_L^3. \end{aligned} \tag{4.16}$$

Here, F, G, H, and I are the areas of summation where three, two, one and no node pairs are of short-range distance, respectively. As shown in Fig. 4.6, one can shift the summations to run from lattice distance $-N/2$ to distance $N/2$ around the focal node at $d = 0$ such that by marking the conditions for short-range connections, finding the respective areas reduces to a geometrical exercise. By defining the lengths $L = (N-1)/2$ and $R = k/2$ one first finds the useful unit of a short-short-long-range area as the triangle $T = (R^2 - R)/2 + R$.

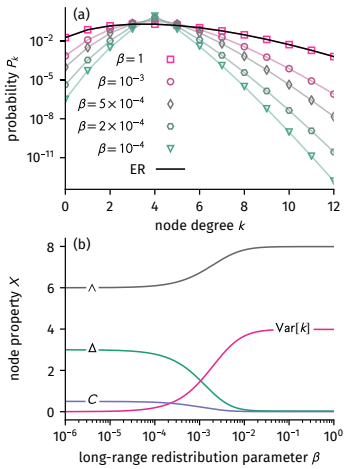


Figure 4.7: Analytic results for **(a)** degree distribution Eq. (4.14) and **(b)**, expected number of two-stars per node $\wedge$ (Eq. (4.15)), expected number of triangles per node $\triangle$ (Eq. (4.16)), clustering coefficient $C = \triangle/\wedge$ and node degree variance Eq. (4.13). While both degree variance and number of two-stars increase with increasing long-range redistribution parameter β, the number of triangles as well as the clustering coefficient decrease. The results shown here were computed for $N = 1001$ and $k = 4$. The figure was taken from [94].

Then, the areas of summation are

$$\begin{aligned} F &= 3(T-R) = (3k/8)(k-2) \\ G &= 3T = (3k/8)(k+2) \\ H &= 2((L-R)R-T) + T + 2(L-R)R + 2((L-1)R-T) \\ &= (k/8)(12N-26-11k) \\ I &= (L-R)^2 - 2((L-1)R-T) - (L-R) + (L-R)^2 - T \\ &= (1/8)\left[5k^2 - k(12N-26) + 4(N^2-3N+2)\right]. \end{aligned}$$

The expected number of triangles decreases with increasing β, as expected and as shown in Fig. 4.7b. Considering Eqs. (4.12), the clustering coefficient is consequently given by

$$C(\beta) = p_S^3 \times \frac{F + G\beta + H\beta^2 + I\beta^3}{(1/2)\mathrm{Var}\,[k] + k(k-1)}. \tag{4.17}$$

In the respective limits one finds

$$\begin{aligned} C(\beta = 0) &= \frac{3(k-2)}{4(k-1)} \\ C(\beta = 1) &= \frac{\sum_{u=2}^{N-1}\sum_{v=u+1}^{N} p^3}{\sum_{u=2}^{N-1}\sum_{v=u+1}^{N} p^2} = p, \end{aligned}$$

which are the expected results for both the k-nearest neighbor lattice (see Eq. (2.6)) as well as the Erdős–Rényi graph (see Eq. 2.7). Further considering Eqs. (4.12) and (4.17) as well as noting that $\mathrm{Var}\,[k]\,(\beta \to 0) = 0$, in the limit of small long-range redistribution one finds

$$\frac{C(\beta \ll 1)}{C(0)} \approx p_S^3 = 1 - 3\beta\frac{N-k-1}{k} + \mathcal{O}(\beta^2) \tag{4.18}$$

which will be of importance for quantifiying the small-world effect in Sec. 5.1.2.

4.3 *Power-Law Small-World (PLSW) Networks*

In the original Kleinberg model (see [31] and Sec. 2.2.4), N nodes are embedded in a low-dimensional space and positioned as a lattice, such that they are connected to their nearest neighbors. Additional long-range links are then added with probability P that decays as a power-law with lattice distance d between two nodes, following $P \propto d^{-\kappa}$, with $\kappa > 0$. This procedure is limited in its application as it constrains the basic topology to low-dimensional lattices and does not account for a constant expected number of generated edges. Furthermore, it does not approach the Erdős–Rényi random graph with $\kappa = 0$, which would be favorable for comparisons. For these network models, the power-law exponent $-\kappa$ is the *structural control parameter* which controls the amount of long-range edges and their distance distribution.

In the following, a generalized variant of Kleinberg's model is considered where all pairs of nodes are connected using a connection

probability function which decays as a power-law with their distance in an embedding space (thus, neighbors on lattice points are not automatically connected). It will be shown that edge distances do not necessarily have to be distances and thus the embedding topology is not of importance to be known. The model is therefore generalized to categorical distances where edges can belong to categories associated with positive integers or real numbers which is subsequently used to define a modified Kleinberg model using a one-dimensional ring lattice as an embedding discrete space.

4.3.1 *General Construction of PLSW Networks with Constant Mean Degree*

In the following, a procedure is described which allows to keep the mean degree fixed when varying a distance-based connection probability $P(r)$. The model initially assumes that nodes are distributed in some arbitrary low-dimensional space. Since nodes will be connected with a distance-based connection probability $P(r)$, it is furthermore required that there exists a metric which allows to compute distances r between nodes, such that the distribution of distances between them is captured by the connection density $f(r)$ where $0 \leq r \leq r_{\max}$ with

$$\int_0^{r_{\max}} dr\, f(r) = 1.$$

The connection density is used to evaluate the node pair density $\mathcal{M}(r)$ at distance r which is needed to evaluate how many potential edges are expected to be found for each absolute distance r such that the connection probability $P(r)$ can be normed to produce a constant mean degree afterwards. The node pair density has to integrate to the total number of node pairs. In a system of N nodes this is equal to the number of edges in a corresponding complete network $m = (1/2)N(N-1)$, such that

$$\int_0^{r_{\max}} dr\, \mathcal{M}(r) = \frac{1}{2}N(N-1)$$

and hence

$$\mathcal{M}(r) = \frac{1}{2}N(N-1)f(r).$$

Now, as per definition of the model, the connection probability $P(r)$ should roughly assume the shape of a power-law $C_\kappa r^{-\kappa}$. Let $\epsilon \ll r_{\max}$ be a minimal radius where nodes within a distance of ϵ will always connect, then the connection probability of a node pair at distance r is defined as

$$P(r) = \begin{cases} 1 & 0 < r \leq \epsilon\,, \\ C_\kappa r^{-\kappa} & \epsilon < r. \end{cases} \tag{4.19}$$

The prefactor C_κ can be found by fixing the expected total number of edges in the network as given by $\langle m \rangle = 2\langle k \rangle / N$ with $\langle k \rangle$ being the

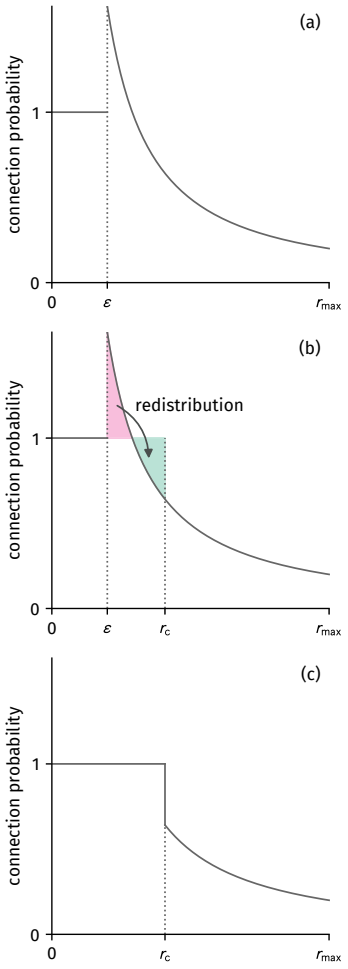


Figure 4.8: How to construct the connection probability such that the expected number of generated edges will be constant for each exponent κ. **(a)** Start with power-law Eq. (4.19). Note that in the shown case the connection probability exceeds $P = 1$ such that this excess probability cannot generate any edges. **(b)** Find the critical distance r_c at which the excess number of edges which cannot be generated will be balanced by redistributing their generation probability to higher distances (see Eq. (4.23)); in this particular example a geometry was chosen where the distance distribution $f(r)$ is uniform. **(c)** The adjusted (final) connection probability.

mean degree. The corresponding integral is

$$\langle m \rangle = \frac{2\langle k \rangle}{N} = \int_0^{r_{\max}} dr\ \mathcal{M}(r)P(r) \tag{4.20}$$
$$= \int_0^{\epsilon} dr\ \mathcal{M}(r) + \int_{\epsilon}^{r_{\max}} dr\ \mathcal{M}(r)C_\kappa r^{-\kappa}$$

which evalutes to

$$C_\kappa = \frac{\frac{\langle k \rangle}{N-1} - \int_0^{\epsilon} dr\ f(r)}{\int_{\epsilon}^{r_{\max}} dr\ f(r)r^{-\kappa}}. \tag{4.21}$$

In many systems and especially for $\kappa \ll 1$ the prefactor C_κ will produce connection probabilities that exceed unity for small distances. This in turn will cause missing edges because those distances r with generation probability $P(r) > 1$ can not contribute more edges than given by the number of node pairs $\mathcal{M}(r)dr$. In order to avoid that edges cannot be constructed due to this circumstance, a connection probability redistribution procedure is introduced as follows.

For all distances where $P(r) > 1$, the total amount of non-generated edges $\mathcal{M}(r)(P(r) - 1)$ is summed up. This excess probability is then redistributed to higher distances until a critical radius r_c is reached, meaning that within this radius, edges are produced from all node pairs (i.e. with probability 1) while outside of this radius edges are produced according to the demanded power-law with exponent $-\kappa$. The connection probability thus changes to

$$P(r) = \begin{cases} 1, & \text{if } 0 < r \leq r_c\,, \\ C_\kappa r^{-\kappa}, & \text{otherwise,} \end{cases} \tag{4.22}$$

where

$$r_c = \begin{cases} \epsilon, & \text{if } C_\kappa \epsilon^{-\kappa} > 1\,, \\ r_{\min}, & \text{otherwise.} \end{cases}$$

The minimal radius $r_{\min}$ is determined by the balance equation

$$\int_{\epsilon}^{r_{\min}} dr\ f(r)\left[C_\kappa r^{-\kappa} - 1\right] = 0 \tag{4.23}$$

with the condition $r_{\min} > \epsilon$, which can be solved numerically with Newton's method for finding zeros and Simpson's rule for computing integrals.[2] This procedure is illustrated in Fig. 4.8.

Generating networks using the connection probability defined above will result in homogeneous networks interpolating between two limit structures. In the first limit $\kappa \to \infty$, the connection probability is

$$\lim_{\kappa \to \infty} P(r) = \begin{cases} 1 & \text{if } r \leq r_c, \\ 0 & \text{otherwise,} \end{cases}$$

[2] Networks sampled from this model within the scope of this work will be generated with $r_{\min}$ as computed with $\epsilon / r_{\max} = 10^{-9}$. If solving Eq. (4.23) numerically failed with this value, ϵ was iteratively doubled until the equation was numerically solvable.

meaning that the generated network will be equal to a random geometric graph (RGG) on the embedding topology with properties $\mathcal{M}(r)$ and $r_{\max}$ (RGGs were introduced in Sec. 2.2.3). In the other limit $\kappa = 0$, it becomes clear that the connection probability becomes distance-independent as $P(r) = C_0 r^0 = C_0$. Furthermore, taking the limit $\epsilon \to 0$ in Eq. (4.21), yields $C_0 = \langle k \rangle / (N-1)$ and thus the connection probability is

$$\lim_{\kappa \to 0, \epsilon \to 0} P(r) = \frac{\langle k \rangle}{N-1} = p_{\mathrm{ER}}.$$

Due to its independence of the distance r, the resulting model is the Erdős–Rényi random graph model with uniform connection probability p_{ER}. Note that as per Eq. (4.20) the distance of generated edges in the sampled network will be distributed according to

$$\mathcal{D}(r) = \frac{2}{N \langle k \rangle} \mathcal{M}(r) P(r). \tag{4.24}$$

The derivation of this procedure assumes that the topology of the embedding space is known. If, instead, simply all $(1/2)N(N-1)$ distances r_{uv} between pairs of nodes are given, the generalized procedure described in App. B.2 may be used, which basically approximates the connection density as

$$f(r) = \frac{2}{N(N-1)} \sum_{u<v} \delta(r - r_{uv})$$

where $\delta(x)$ is Dirac's delta function. Since the topology of the embedding space can therefore be unknown it is important to note that in this more general case, r does not need to be a 'distance' anymore: if only values of r_{uv} between node u and v are known, they do not have to meet the triangle inequality. Instead, these values could stem from semimetrics but also represent rank values or other positive real numbers, i.e. general edge weights. For the considerations above it was further assumed that distances are positive real numbers and hence distributed according to a probability *density* function. However, there exist cases where distances are integers (e.g. for lattices) or not even distances but just categorizations of edges and thus follow a probability *mass* function. Special cases are the modified Watts–Strogatz model where only two distances exist or the SSMH model where nodes can be of hierarchical distance ℓ. For these cases a generalized PLSW connection procedure can be defined as described in App. B.2.2.

4.3.2 *Application to 2-Torus and Square*

As a first, simple application, the unit square $[0,1)^2$ with periodic boundary conditions (i.e. a 2-torus) will be discussed as an embedding space in the following. On this space N nodes are distributed with uniform probability (i.e. their positions x and y follow the same uniform random distribution $x, y \sim \mathcal{U}(0,1)$), yielding a certain pairwise distance distribution $f(r)$ which is derived in App. B.2 and given by Eq. (B.1).

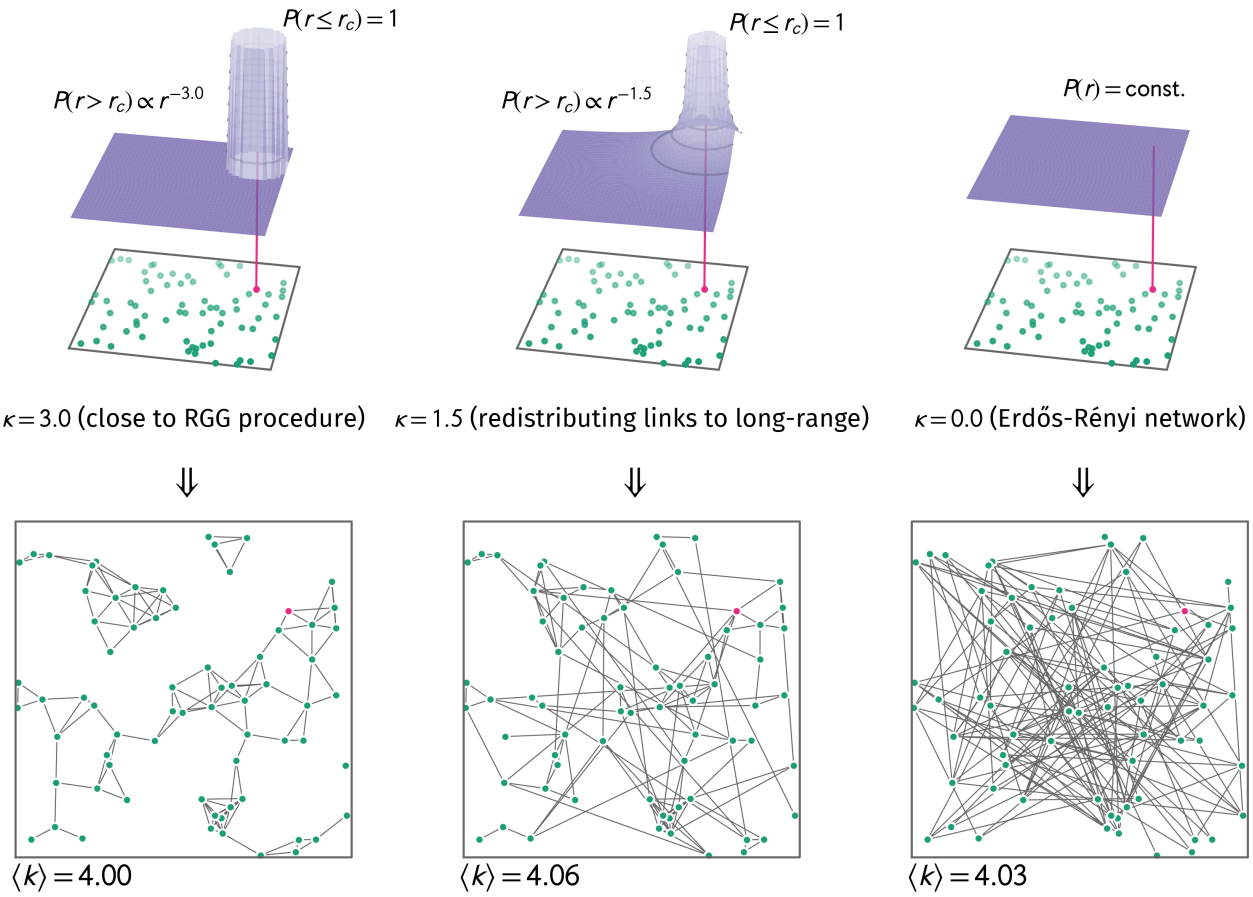


Figure 4.9: Sampling procedure for nodes embedded in $[0,1)^2$ without periodic boundary conditions for parameters for $N = 70$, $\langle k \rangle = 4$, $\epsilon = 10^{-9}$ and varying power-law exponent κ. **(Top row)** Illustration of the distance-varying connection probability function for increasing power-law exponent $-\kappa$ for a single node (pink) as defined by Eqs. (4.22), (4.23), and (B.2). **(Bottom row)** Corresponding generated networks. For connection probability strongly decreasing with distance ($-\kappa \ll 0$), the sampling procedure essentially generates a random geometric graph. With increasing power-law exponent $-\kappa$ more long-range links are produced until the network model equals an Erdős–Rényi random graph at $\kappa = 0$.

Similarly, the distance distribution of two random points in the unit square can be found when *no* periodic boundary conditions are used (see App. B.2). For the latter embedding geometry, example networks and the connection probability function defined in Sec. 4.3.1 are shown in Fig. 4.9. As one can see for $-\kappa \ll 0$ the procedure generates structures which are essentially equal to samples from the random geometric graph model. With increasing power-law exponent $-\kappa$, the connection probability broadens towards higher distances which introduces more long-range links until each edge is equally probable at $\kappa = 0$ to obtain the Erdős–Rényi model. As a numerical validation of the considerations in Sec. 4.3.1, networks on both topologies were generated using a simple sampling procedure as described in App. B.2. The numerically found edge distance distributions are indeed in close agreement with Eq. (4.24).

The generated networks should reproduce the small-world effect as described in Sec. 2.2.4, meaning that the average shortest path length should decline faster than the mean clustering coefficient with increasing power-law exponent $-\kappa$. Additionally, the clustering coefficient is expected to interpolate between the results for the random geometric graph Eq. (2.11) and the Erdős–Rényi graph Eq. (2.7). To test these hypotheses, the corresponding network properties were

computed as averages over 50 network samples for each parameter combination of $N = 1000$, $\langle k \rangle = 10$, $\epsilon = 10^{-9}$ and increasing power-law exponent $-3 \leq -\kappa \leq -1/2$. Both the unit square and the 2-torus were tested as underlying topologies. Note that the average shortest path length $\langle s \rangle$ was only computed on the largest connected component of a network which was always of size $\approx N$.

The numerical results are shown in Fig. 4.10. As expected, the small-world effect is displayed with increasing power-law exponent, meaning that the average shortest path length decreases faster than the clustering coefficient with increasing introduction of long-range links. Note that the expected theoretical limit results of the clustering coefficient are reached (with higher accuracy for the 2-torus since the periodic boundary conditions prevent geometrical deviations). The mean degree $\langle k \rangle$ remains constant (bar rather small fluctuations) for varying control parameter, as demanded per construction. Surprisingly, both the average number of two-stars per node as well as the degree variance remain constant (small deviations for the square-topolgy without periodic boundaries, respectively). However, since in the RGG limit the probability to connect to k nodes follows a binomial distribution with the success probability being equal to the disk area of radius r_c around a focal node, the variance might be indeed expected to be equal to that of the Erdős–Rényi model, hinting at the fact that it simply stays constant.

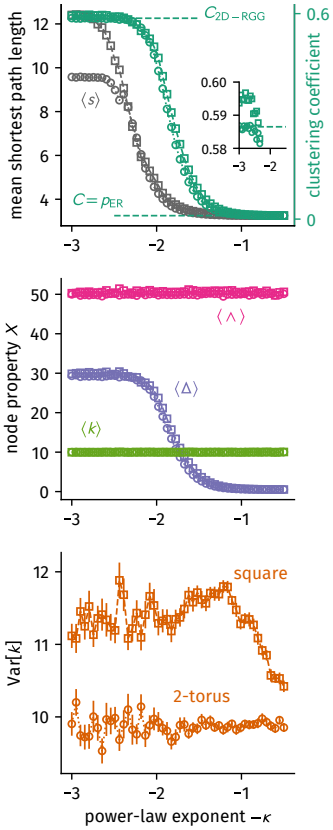


Figure 4.10: Network properties of PLSW small-world networks with underlying topology of the unit square (curves marked with □) and the 2-torus (curves marked with ○). Each point is the mean over 50 independent realizations for $N = 1000$, $\langle k \rangle = 10$, $\epsilon = 10^{-9}$, and increasing power-law exponent $-3 \leq -\kappa \leq -1/2$. For the mean shortest path length $\langle s \rangle$, the largest connected component of the generated network was used. **(Top)** The mean shortest path length decreases faster than the clustering coefficient, thus displaying the small-world effect. The expected theoretical values for the clustering coefficient of both limit structures are reached. Inset: The results for the 2-torus better match the theoretical value of Eq. (2.11) than the results for the unit square since for the latter case and finite N the hard boundary skews the overlap of circles for boundary-near nodes. **(Middle)** Mean degree and number of two-stars per node remain constant while the number of triangles per node decreases. **(Bottom)** The degree variance of the networks generated with the 2-torus topology remains constant.

4.3.3 *Application to One-Dimensional Lattice PLSW Networks*

In this subsection, a PLSW network model is introduced which is conceptually comparable to the Kleinberg model in Sec. 2.2.4. Parts of the text given below are close to the published article [93].

The model can be described as follows. Beginning similarly as with the modified Watts–Strogatz (WS) model introduced in Sec. 4.2.1, given two distinct nodes u and v with indices $0 \leq u \leq N-1$ and $0 \leq v \leq N-1$ positioned equidistantly on a ring, the shortest distance between them is

$$d(u,v) = \min(|v-u|,\ N-|v-u|).$$

In contrast to the modified WS model, the connection probability between any two nodes is then defined by a power-law as

$$P_{d(u,v)} = p_0 |d|^{\mu-1}. \tag{4.25}$$

Here, $-\infty < \mu \leq 1$ is the respective structural control parameter controlling the amout of long-range edges (larger μ values imply higher long-range connection probability).[3] Note that again, this lattice PLSW model will produce a problem for large negative values of μ, for which the connection probability between close neighbors will exceed unity. As was argued above, this excess probability can be redistributed to the nearest neighbors until one runs out of excess probability, as illustrated in Fig. 4.12, essentially replacing Eq. (4.25)

[3] Note that requiring a constant mean degree, the normalizing constant evaluates to $p_0 = \langle k \rangle \Big/ \left[2\sum_{d=1}^{\lfloor N/2 \rfloor} d^{\mu-1} + \mathrm{mod}(N,2)(\lfloor N/2 \rfloor + 1)^{\mu-1}\right]$, where $\mathrm{mod}(N,2)$ denotes the remainder of the integer division $N/2$ and $\lfloor \cdot \rfloor$ is the floor function.

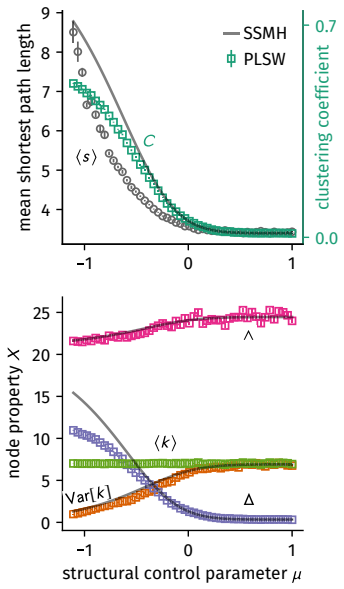


Figure 4.11: Properties of the one-dimensional lattice PLSW network model for increasing control parameter μ, here with parameters $N = 8^3$, and $\langle k \rangle = 7$. The observables displayed here were obtained as an average over 10 independently sampled networks for each value of μ. **(Top)** The mean shortest path length $\langle s \rangle$ was computed on the largest component of the sampled networks and decreases faster than the clustering coefficient $\langle C \rangle$, which is compared to the corresponding SSMH clustering coefficient given via Eq. (4.11) and using the map of the control parameters Eq. (4.27). The SSMH clustering coefficient estimates this PLSW clustering coefficient well for larger values of the control parameters but deviates when approaching the regular structures (since those are fundamentally different). **(Bottom)** The mean number of triangles per node (compared to the SSMH result Eq. (4.9)), the mean number of two-stars per node (compared to the SSMH result Eq. (4.10)), the node degree variance (compared to the SSMH result Eq. (4.8)), and the mean degree $\langle k \rangle$. Note that none of the observables change significantly after leaving the weak modular hierarchically clustered phase at $\mu = 0$.

with Eq. (B.3). This results in producing a $\langle k \rangle$-nearest neighbor lattice for $\mu \ll 0$, similar to all variants of the Watts–Strogatz small-world model. For a sampling algorithm, loop over all node pairs $0 \leq u \leq N-2$ and $u+1 \leq v \leq N-1$, compute the distance $d(u,v)$ and connect this pair according to this modified pmf. This algorithm is of order $\mathcal{O}(N^2)$.

Samples from this network model show similar structural behavior as the models discussed previously, as can be seen in Fig. 4.11. Again, with increasing structural control parameter, the mean shortest path length decays faster than the clustering coefficient and the degree variance increases. Example networks for increasing structural control parameter are displayed in Fig. 4.14.

4.4 *Structural Relation of Small-World Models*

In this section, three of the models introduced in this chapter will be shown to be structurally similar by relating their respective structural control parameters and notion of distance. The three models are the self-similar modular hierarchical (SSMH) small-world model, the one-dimensional lattice PLSW network model and the modified Watts–Strogatz model. It will be shown that all of these models appear to produce modular hierarchical structure.

4.4.1 *Mapping SSMH Networks to One-Dimensional Lattice PLSW Networks*

Parts of the text given below are close to the published article [93]. In order to show that SSMH and one-dimensional lattice PLSW are conceptually similar, the hierarchical distance between nodes will be mapped to the lattice distance on a ring (i.e. to the defining distance in the latter model). The hierarchical distance from a focal node at layer $\ell = 0$ may first be approximated as a continuous function. Then, the connection probability density in layer space as per Eq. (4.2) is given by

$$p(\ell) \propto \frac{\xi^\ell}{B^\ell}. \tag{4.26}$$

The total number of nodes reachable by the focal node up to hierarchical distance ℓ is given as $n(\ell) = B^\ell$ at thus the infinitesimal number of potential neighbors at ℓ is

$$dn = \frac{dn}{d\ell} d\ell = B^\ell \log B \; d\ell.$$

The number of connections the focal node has to layer ℓ is therefore

$$\begin{aligned} dk &= p(\ell) \; dn \\ &= \xi^\ell p_0 \log B \; d\ell. \end{aligned}$$

One may define the connection density (connections per layer) in layer space as

$$dk = \frac{dk}{d\ell} d\ell,$$

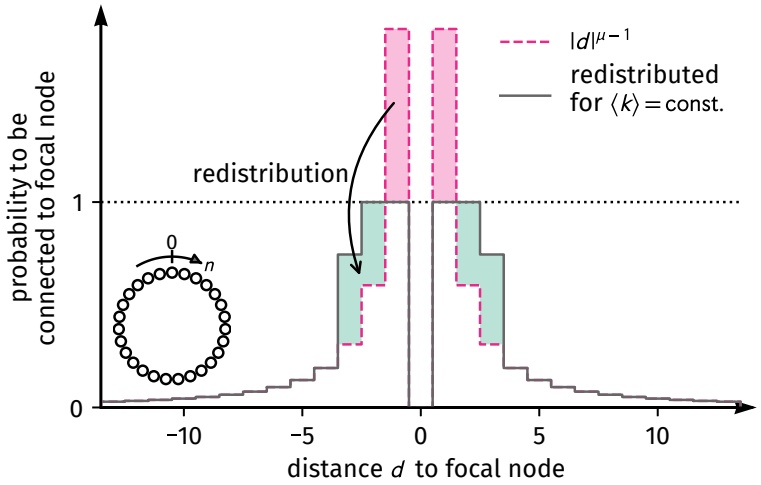


Figure 4.12: Method for redistributing the probability mass function of the one-dimensional lattice PLSW network model such that it does not exceed unity in numerical calculations. The original connection probability (red dashed line) is modified by redistributing the excess probability to its nearest neighbors. The figure was reproduced from [93].

so one finally obtains

$$\frac{dk}{d\ell} = \xi^{\ell} p_0 \log B \equiv p(\ell).$$

Now, with $p(y) = \frac{dy}{dx} p(x)$ and $\ell(n) = \log n / \log B$ one finds

$$p(n) = p_0 n^{\mu-1}$$

which is the distance-based power-law connection probability in a one-dimensional lattice and thus relates the two respective structural control parameters as

$$\mu = \frac{\log \xi}{\log B}. \tag{4.27}$$

This result can be interpreted as follows. The SSMH connection probability function corresponds approximately to a discretized version of the power-law that is associated to the PLSW connection probability, however, target nodes are grouped in exponentially growing batches (cf. Fig. 4.13). In this approximation, the additional embedded nature of the SSMH network structure is lost. It thus becomes clear that SSMH networks are a special case of the general categorized PLSW model of Sec. B.2.2.

One important structural result of this mapping is that for the SSMH structural control parameter of value $\xi = 1$, where all hierarchical clustering is lost as per Eq. (4.5), the corresponding PLSW control parameter is $\mu = 0$, resulting in the connection probability function $P(n) \propto n^{-1}$ which is the critical connection probability function at which locally informed search processes become optimal (see Sec. 2.2.4). A comparison of example networks for equal structural control parameters are shown in Fig. 4.14.

4.4.2 *Mapping modified Watts–Strogatz Networks to One-Dimensional Lattice PLSW Networks*

In order to relate the structural control parameters of the modified Watts–Strogatz model β and the one-dimensional lattice PLSW

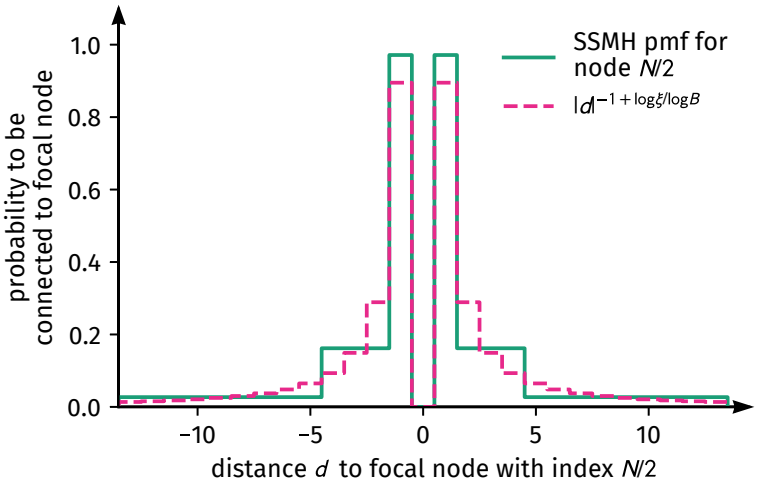


Figure 4.13: Comparison of the connection probability functions associated with SSMH networks and with one-dimensional lattice PLSW networks. Parameter values chosen here are number of submodule per module $B = 3$, hierarchical layers $L = 3$, mean degree $\langle k \rangle = 3.4$ and structural control parameter $\xi = 0.5$ which relates to $\mu \approx -0.63$. The SSMH connection probability function can be interpreted to group nodes of similar distance to the focal node to exponentially growing batches of equal connection probability. The figure was reproduced from [93].

model μ, one recognizes first that there are three important structural conditions for the control parameters. As per definition of these small-world models, the two important limits are reached when (a) $\mu \to -\infty$ and $\beta \to 0$, which is when both models become structurally equivalent to a k-nearest neighbor lattice, and (b) for $\mu = 1$ and $\beta = 1$ which are the values where all network models become equivalent to the Erdős–Rényi model. One formulaic way to ensure that these values are reached is

$$\log \beta = A\mu + B. \tag{4.28}$$

This equation can be motivated as follows. Since both models are structurally equivalent for small values of β (large negative values of μ), it seems to be appropriate to demand that the short-range degree k_S of both models corresponds when relating their control parameters. For the modified Watts–Strogatz model one finds $k_S = kp_S = k/[1 + \beta(N - 1 - k)/k]$. For the PLSW model, one may approximate the discrete probability connection function as a continuous function to obtain the short-range degree

$$\begin{aligned} k_S &\approx k \int_1^{k/2} x^{\mu-1} dx \Big/ \int_1^{(N-1)/2} x^{\mu-1} dx \\ &= k \frac{(k/2)^\mu - 1}{((N-1)/2)^\mu - 1} \end{aligned}$$

and thus

$$\beta \approx \frac{k}{N-1-k} \left[\frac{((N-1)/2)^\mu - 1}{(k/2)^\mu - 1} - 1 \right],$$

giving justification to assume an approximate form of Eq. (4.28) for relating the control parameters. Now, since this introduced relation formula Eq. (4.28) already meets two important structural limits, a third one is considered to find the values of the parameters A and B in the following.

As discussed in Sec. 4.5, the SSMH structural control parameter value at which all hierarchical structure is lost is given as $\xi = 1$ (or

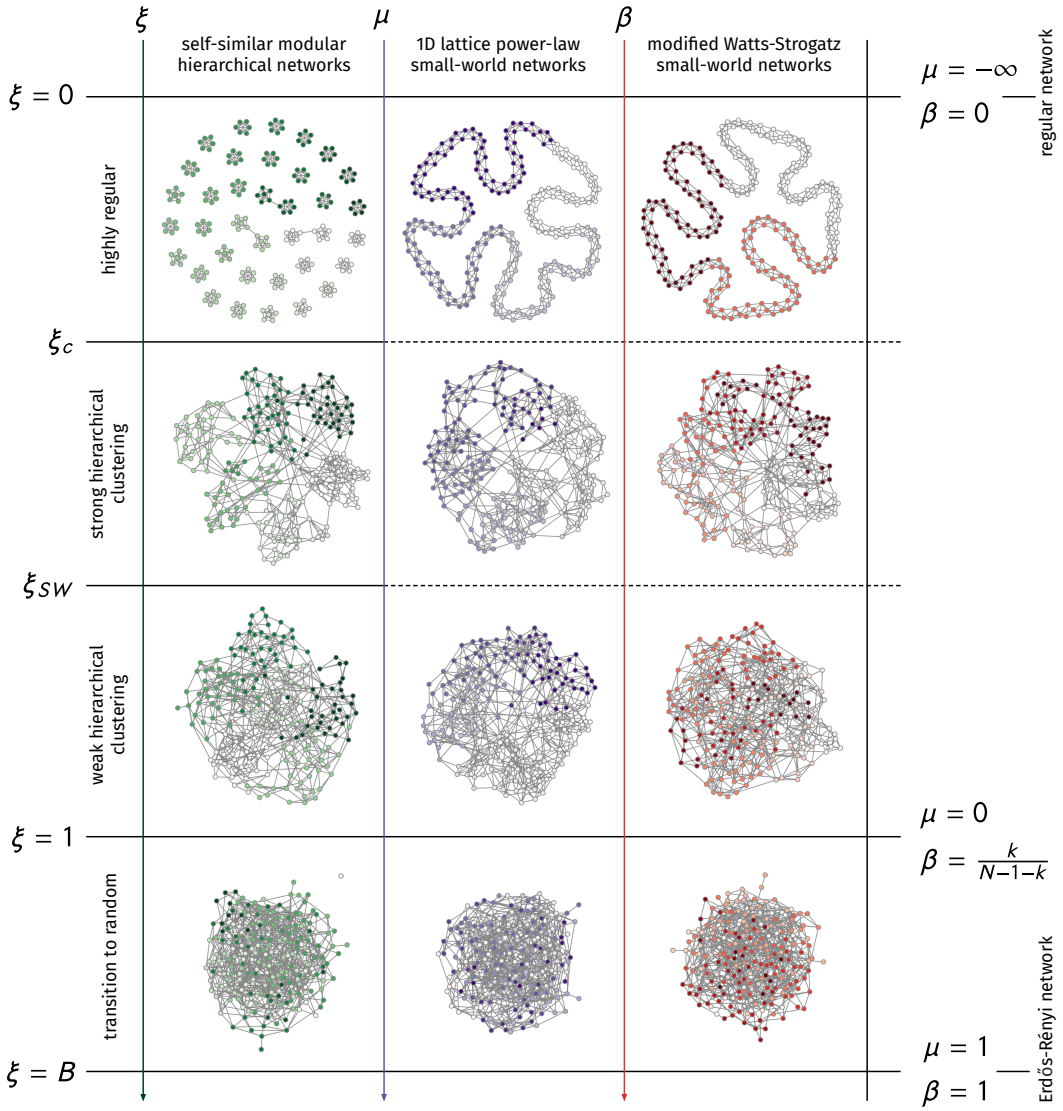


Figure 4.14: Topological phases of three small-world models as a function of their respective structural control parameters ξ, μ, and β as mapped to each other by Eqs. (4.27) and (4.29). Chosen parameters are mean degree $\langle k \rangle = 6$, base number of submodules $B = 6$, and number of hierarchical layers $L = 3$ (number of nodes $N = 6^3$, respectively). For small values of the structural control parameters the sampled networks closely resemble their base structures. With increasing long-range connection probability, all of the network samples seem to display sort of a modular hierarchical structure. While in the SSMH model (green, left) this modular hierarchical structure is given by design, the one-dimensional lattice PLSW (middle, violet) and modified Watts–Strogatz models (right, orange) would display modular hierarchical clustering only as a result from noise (see Sec. 4.4.3). The network figures were generated using a custom interactive visualization tool made by the author which is available online at `https://bit.ly/2TrhUfs`.

$\mu = 0$, respectively). This is the structural point at which a random node has, on average, as many connections to nodes in short range as it has to nodes in each of the categories of greater hierarchical distance. For the modified Watts–Strogatz model, there are only two categories: short-range and long-range. Hence, it seems plausible to connect the control parameters in such a way that at $\xi = 1$ (or $\mu = 0$, respectively), the short-range degree is equal to the long-range degree, which as per Eq. (4.12) yields

$$\begin{aligned} k p_S &= (N - 1 - k)\beta p_S \\ \beta &= k/(N - 1 - k). \end{aligned}$$

In order for the mapping to reach all structural limits as discussed as well as meet Eq. (4.28), a possible choice is hence

$$\beta = \frac{k}{N - 1 - k} \left(\frac{N - 1 - k}{k} \right)^{\mu}. \tag{4.29}$$

Since usually $N - 1 - k > k$, a simple check finds the three demanded

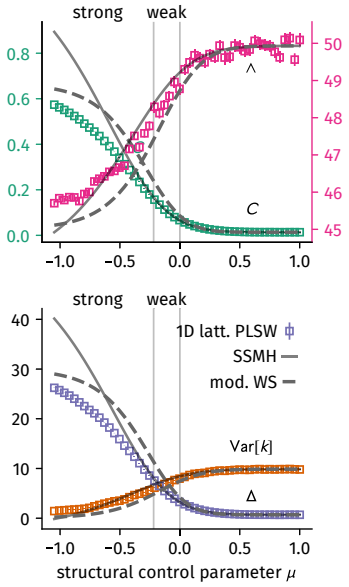


Figure 4.15: Node properties of the SSMH, the modified WS and the one-dimensional lattice PLSW network model for increasing control parameter μ, here with parameters $B = 8$, $L = 3$ (number of nodes $N = 9^3$), and $\langle k \rangle = 10$. The observables displayed as markers were obtained as an average over 100 independently sampled networks for each value of μ. The solid lines represent the theoretical results of the SSMH model and the dashed lines represent the theoretic results of the modified Watts–Strogatz model. Note that none of the observables change significantly after leaving the weak modular hierarchically clustered phase at $\mu = 0$.

values $\beta(\mu \to -\infty) = 0$, $\beta(\mu = 0) = k/(N-1-k)$, and $\beta(\mu = 1) = 1$. A comparison of example networks for equal structural control parameters is shown in Fig. 4.14 and a comparison of network properties is presented in Fig. 4.15. Comparing the sample networks with the naked eye hints at the existence the correspondence of these models, which seems plausible comparing the network properties, as well. This will be further discussed below.

4.4.3 *All Small-World Network Models Appear to be of Modular Hierarchical Structure*

In the former discussion within this chapter, the claim that the models defined above are structurally equivalent was made a handful of times, supported by the analysis of several network or node properties which indeed behave similarly when varying their respective control parameters. However, one might wonder how similar the networks are in the literal sense of their modular hierarchical structure. In the following, the modular hierarchical network partition algorithm *Infomap* will be used to find optimal modular hierarchical partition of networks generated from the models defined above; the algorithm is described in Sec. 2.1.7 and App. A. It finds such a partition by minimizing the minimum description length of an ergodic random walk on a network. The found modular hierarchical partition will be quantified by the height L of its tree where leaves represent nodes and internal nodes represent groupings of nodes at different coarse grained levels, as well as the mean number $\langle B \rangle$ and standard deviation $\mathrm{Std}[B]$ of submodules per internal node (see Fig. 2.8).

For the analysis of the modular hierarchical network partitioning of the model networks, network samples were generated for parameters $B = 8$, $L = 3$ and $\langle k \rangle = 8$ for the SSMH, the one-dimensional lattice PLSW, the modified Watts–Strogatz and the 2-torus PLSW model, and their respective structural control parameters were varied. For every parameter combination, 100 networks were sampled from their respective models. Subsequently, *Infomap* was used with default parameters to find an optimal modular hierarchical network partition. The results are shown in Fig. 4.16.

As one can see, all of the networks are found to be of modular hierarchical structure within the corresponding phases of the structural control parameters $\mu < 0$ and $\kappa < -2$, respectively, finding a number of hierarchical layers $L \geq 3$ to best describe their network structures. Similarly, the mean number of submodules is close to $\langle B \rangle = 10$ ($\langle B \rangle = 8$ for the SSMH model, as expected from the construction condition $B = 8$). Remarkably, for the SSMH model, *Infomap* recovers an underlying hierarchical tree which is almost exactly structurally equal to the original underlying hierarchical tree with the standard deviation of number of submodules per internal node being close to zero. For the other models, this standard deviation is greater but approximately half of the mean. A nonzero standard deviation is

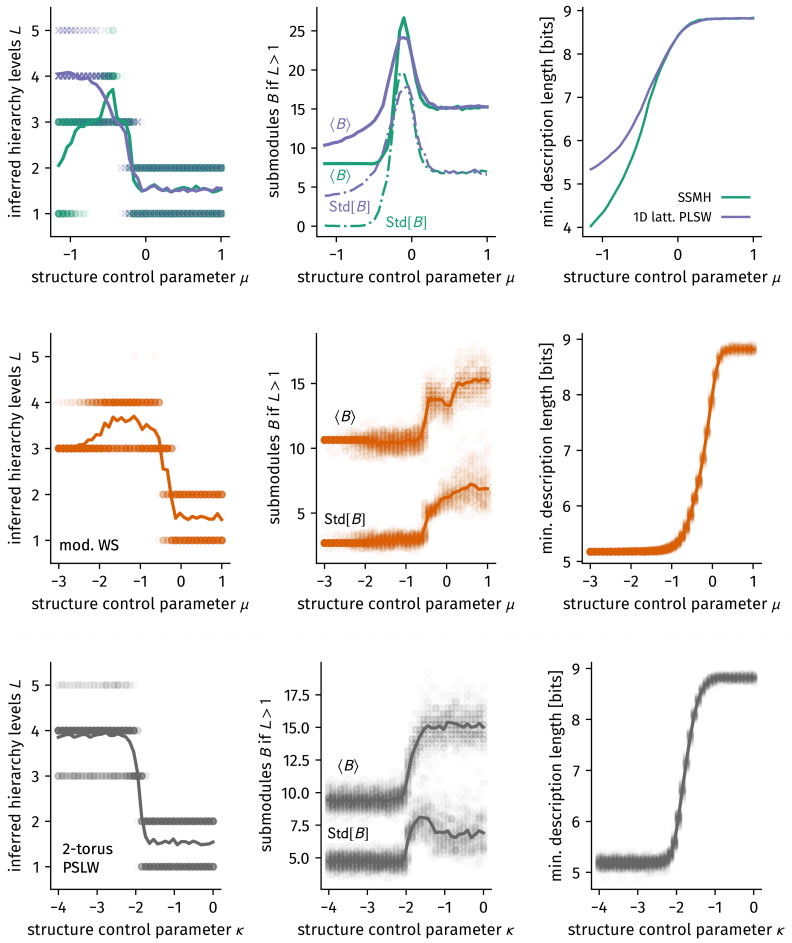


Figure 4.16: Results of the modular hierarchical partition detection using *Infomap* for $B = 8$, $L = 3$ ($N = 8^3$), and $\langle k \rangle = 8$. For each parameter combination, 100 networks were sampled and subsequently analyzed using *Infomap*. Markers represent results for single networks and solid lines are averages over all samples. Dashed lines in the top row are averages of the standard deviation of number of submodules per network sample. **(Top row)** The SSMH (green) and 1D lattice PLSW model (violet). **(Middle row)** The modified Watts–Strogatz model. **(Bottom row)** The 2-torus PLSW model. **(Left column)** The number of measured hierarchy levels L. **(Middle column)** Mean and standard deviation of number of submodules B_I per internal node I. **(Right column)** Minimum description length of the found network partition.

to be expected since there is no underlying hierarchical tree for the other models and all correspondences to a hidden hierarchical structure are due to noise: By chance, a certain amount of potential edges in the same area will not be sampled while others are, producing randomly appearing blocks of ones in the adjacency matrix which are then found by the extraction algorithm. All network structures seem to no longer be able to be resolved when their structural control parameters come close to some threshold which is determined by the exponent of the power-law connection probability being equal to their respective embedding dimension (hence, $\kappa = -2$ for the 2-torus PLSW model and $\mu = 0$ as $P_d \propto d^{\mu-1}$ for the rest). This is the point where an average node has as many connections in short-range as it has to any distance in the long-range category (cf. Eq. 4.5). One might expect this effect, as an algorithm cannot resolve a formerly defined group membership if a node does not have more connections to nodes of this group than to nodes in other groups [29]. The modified Watts–Strogatz model seems to lose its resolvable structure a bit sooner than the SSMH and the one-dimensional lattice PLSW models. Hence, one could interprete this model as being somewhat less modular hierarchical. All of the models which are not modular hierarchical by design seem to converge to a number of ≈ 5

bits minimum description length, compared to a compressibility of ≈ 9 bits for the Erdős–Rényi case,[4] whereas the SSMH model approaches ≈ 4 bits compressibility. The SSMH model's compressibility is smaller than the other models' compressibility because the SSMH model is truly generated from a hierarchical tree, which can be found, whereas the others are not and modular hierarchical structure is produced by chance, as argued above.

To sum up, all the models seem to generate networks which are interpretable as modular hierarchical, at least for a random walker spending time in different regions of the network as modeled by *Infomap*.

[4] For the ER network, Infomap finds a one-level solution where every node is its own module. This means that the per-step minimum description length becomes equal to the Shannon entropy Eq. A.1 with ergodic node visiting probabilities $p_i = k_i / N \langle k \rangle$. Hence, the compressibility can be computed as $H_{\mathrm{ER}} = -(N \langle k \rangle)^{-1} \sum_{i=1}^{N} k_i \log_2(k_i / N \langle k \rangle)$. Using the binomial degree distribution of the ER model and noting that $k_i > 0$, since only the largest component has been used, the minimum description length can be estimated as $H_{\mathrm{ER}} \approx -\left[\langle k \rangle^{-1} \langle k \log_2 k \rangle_{k>0} - \log_2(N \langle k \rangle)\right]$ which can be numerically computed to find $H_{\mathrm{ER}} \approx 8.9$ for the parameters chosen here.

4.5 Summary and Discussion

In this chapter, a modular hierarchical network model was introduced, based on a self-similar hierarchical partition tree. This network model generates modular hierarchically structured networks with a constant mean degree. A single structural control parameter controls the number of long-range edges per node and can be used to interpolate between a structure where all edges are short-ranged (only in modules in the lowest hierarchy module) and the Erdős–Rényi random network model. Clustering coefficient and node degree variance were computed analytically as functions of the structural control parameter. It was shown that different values of the control parameter correspond to different modular hierarchical phases where (i) an average node has more connections to nodes within its lowest-level hierarchy cluster than to nodes of its higher-level clusters combined called the *strong modular hierarchical phase*, (ii) an average node has more connections to nodes within its lowest-level hierarchy cluster than to nodes of each of its higher-level clusters, called the *weak modular hierarchical phase* and (iii) the network is already close to a random network.

As this network model is similar to two other existing network models, namely the Watts–Strogatz and the Kleinberg model, modified versions of these models were introduced. Their modified definitions will allow for a more thorough analysis concerning dynamics as they both keep their mean degree fixed and approach the Erdős–Rényi model in the random limit, thus enabling direct comparisons to a known network model and networks of similar density.

In the original Kleinberg small-world model, nodes are connected based on a connection probability decaying as a power-law with their distance. This power-law procedure was generalized to find that it can be applied to arbitrary embedding geometries as well as arbitrary edge weights which do not have to be distances.

On this basis, an additional model was devised where nodes are uniformly distributed on a 2-torus and subsequently connected with a probability proportional to a power-law that decays with their distance. Contrary to the other models devised in this chapter, this model does not only have a constant mean degree but a constant

node degree variance, as well—while other properties remain similar to those of the other small-world models.

After the definition of the models, it was shown how they correspond to each other by finding equations mapping their respective structural control parameters. It was subsequently found that their network properties are similar when varying their structural control parameters and that a modular hierarchical network partition algorithm finds similar network partitions for networks sampled from all the models defined in this chapter.

Therefore, it has to be assumed that real-world networks which were found to be of modular hierarchical structure do not have to be assumed to be of strict modular hierarchical (block) structure. Instead, it seems likely that a considerable amount of these networks can be described as embedded in an abstract low-dimensional space and nodes being connected with a distance-based power-law connection probability. Indeed, similar considerations about different kinds of community structures have been made, inferring a distance-based community partition with information theoretic methods [28].

Based on the results of this chapter, an additional simple method to test how "modular hierarchical" a network really is might be given as follows. Find the minimum description length of the network. Generate similar networks from the SSMH model where B and L are chosen close to the values of the extracted modular hierarchical tree and scan the modular hierarchical structure parameter ξ. Further generate similar networks from one of the PLSW models and vary their respective modular hierarchical structure parameters. Finding the model which has similar compressibility to the original network might give an indication which of the underlying network models are more likely to be responsible to have generated the network.

Note that structures of power-law small-world networks can be easily interpreted to yield overlapping modules, since there are no underlying 'true' communities. Such structures are also related to modular hierarchical communities of links [71]. Future studies might model these modular hierarchical link partitions by similar models as the ones introduced in this chapter in order to study the existence of a link-centric small-world effect.

Such applications to investigate real-world networks for their modular hierarchical or small-world structure, are, however, left for future research.

Ultimately, the models introduced in this chapter are ideal candidates to study how the outcome of basic dynamics is influenced by (i) the introduction of long-range connections, (ii) the difference of small-world embedding topologies, and (iii) variation of node degree variance, all while (iv) the mean degree remains constant[5] and (v) well-investigated structures are approached in the two structural limits. Therefore, in a first application, these models will be used to test the hypothesis of modular hierarchical networks being of optimal topology for fast random walk processes in the following chapters.

[5] as is important to gauge the outcome of dynamics (see Ch. 3)

5
Passage Time Statistics for Locally Clustered Networks

Modular hierarchical networks are ubiquitous in the real world. In this chapter the hypothesis is tested that such topologies provide an 'optimal structure' for certain dynamic processes as a potential explanation for their omnipresence.[1] As argued before, one of the simplest dynamic processes is diffusion as modeled by random walks (see Sec. 3.1). Indeed, it has been found that locally informed searches are minimal on a modular hierarchical network model (see Sec. 2.2.4 and [13, 31]). Since random walks are a reasonable model for random searches, their temporal observables place a natural upper bound on the corresponding observables of informed searches. Hence, this chapter will evaluate the potential of modular hierarchical networks to provide an optimal structure for random searches, using the network models introduced in Ch. 4.

[1] Naturally, what 'optimal' is depends on the application, but in the following it will refer to minimal temporal observables of random walks.

It will be found that an average medium approximation can be used to analytically estimate the shape of the pair-averaged first passage time (FPT, see Sec. 3.1.5) as a function of a model's respective structural control parameter. This function does not reveal the existence of an optimal structure for fast random searches. A subsequent analysis on actual network realizations will reveal that three of the four network models discussed in Ch. 4 display a minimum in the pair-averaged FPT. This effect is consequently explained by deriving a heuristic lower bound of the pair-averaged FPT for general locally clustered networks, which shows to be composed of a node degree variance distribution as well as a clustering coefficient contribution. As one of them decreases while the other increases with increasing structural control parameters for said three models, they combine to build a minimum. A fourth modular hierarchical network model with constant node degree variance is consequently shown to not possess an optimal structure for random searches. It is therefore argued that modular hierarchical structures do not inherently possess a minimal pair-averaged FPT but that the emergence of this minimum is a pure consequence of network structures with decreasing clustering and increasing node degree variance. Large parts of this chapter are close to the published articles [93, 94].

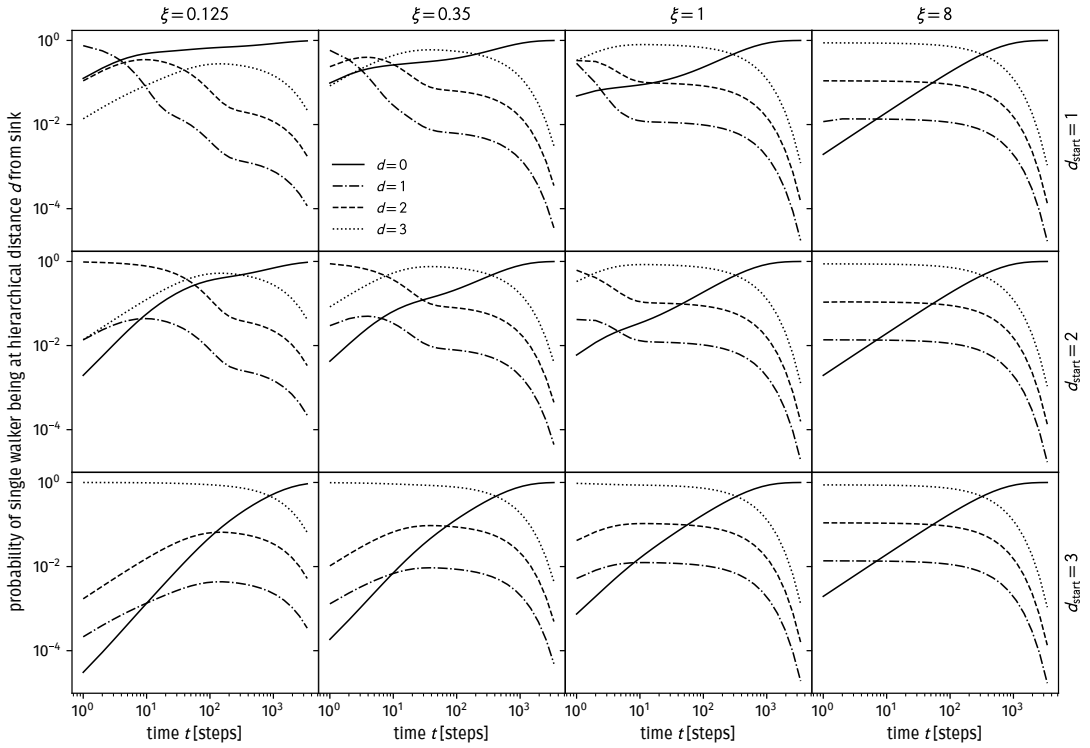


Figure 5.1: Time evolution of the random walker density at hierarchical distance d from the sink when the walk started on a single node at hierarchical distance d_{start} for the average medium approximation as per Eq. (5.3). Parameters used here where $B = 8$ and $L = 3$. Note that the initial condition Eq. (5.2) is not displayed as the plots begin at $t = 1$.

5.1 *Average Medium Approximation*

This section is close to the text first published in [93]. In the context of random walks on networks, an *average medium approximation* (AMA) can be used to solve diffusion problems by approximating the network topology by an 'average structure'.

Here, this AMA is given by connecting all pairs of nodes (u, v) with a link with weight P_{vu}, equal to the probability of connecting both nodes in the original random network. In order to do this, P_{vu} has to be normalized so that, from each node, the total probability to jump to any other node in one time step is equal to one, that is, $\sum_v P_{vu} = 1$. For example, given an Erdős-Rényi network with N nodes, this AMA would connect each node to every other node, except for itself, through a link with weight $P_{vu} = p = 1/(N-1)$ and hence the mean FPT (MFPT) between any pair of nodes would be equal to the MFPT in a complete graph $\tau_{vu} = N - 1$, a result consistent with the large mean-degree limit of more sophisticated AMAs [103]. In this subsection, the FPT statistics and pair-averaged FPT of the SSMH model introduced in Sec. 4.1 will be evaluated. Furthermore, the mixing time of the modified Watt–Strogatz model introduced in Sec 4.2 will be computed analytically and used to illustrate the small-world effect without the use of the average shortest path length. Both observables are shown to monotonically decrease with increasing structural control parameters ξ and β.

5.1.1 SSMH First Passage Time Distribution and Pair-Averaged First Passage Time

Time Evolution of Walker Distribution

Following Eq. (4.3), the probability of two nodes to be connected will be denoted as $\tilde{P}_\ell$. In the context of this AMA, this means that every node is connected to every other node but edges are weighted with probability $P_\ell = \tilde{P}_\ell / \langle k \rangle$. On this averaged network, a random walk with a sink at an arbitrary node v is investigated. When a walker is positioned at node u, the probability to jump to node v at hierarchical distance $d(u,v) \equiv d$ is

$$P_d = \frac{1}{B-1}\left(\frac{1-\xi}{1-\xi^L}\right)\left(\frac{\xi}{B}\right)^{d-1},$$

as reasoned above. The transition matrix from a walker being at distance d from the target to being at new distance $0 \leq d' \leq L$ from the target is derived in App. B.3.1 to be

$$W_{d'd} = \begin{cases} \frac{1}{(B-1)}\frac{1-\xi}{1-\xi^L}\left(\frac{\xi}{B}\right)^{d-1}, & d' = 0 \ \wedge \ d > 0 \\ \frac{B^{d'-1}}{B^{d-1}}\xi^{d-1}\frac{1-\xi}{1-\xi^L}, & d' < d \ \wedge \ d > 0 \\ \frac{1-\xi^{d-1}}{1-\xi^L} + \frac{1-\xi}{1-\xi^L}\xi^{d-1}\frac{B-2}{B-1}, & d' = d \ \wedge \ d > 0 \\ \frac{1-\xi}{1-\xi^L}\xi^{d'-1}, & d' > d \ \wedge \ d > 0 \\ 0, & d = 0 \ \wedge \ d = 0 \\ 1, & d' = 0 \ \wedge \ d = 0. \end{cases} \tag{5.1}$$

Let the $(L+1)$-dimensional vector $\boldsymbol{p}^{(d_{\text{start}})}(t)$ be the probability to find a single walker at distance d when a random walker started at distance $d_{\text{start}} \equiv d_s$ from the target.Let a discrete-time random walk begin with initial conditions given by the vector

$$p_d^{(d_{\text{start}})}(0) = \delta_{d d_{\text{start}}}, \tag{5.2}$$

i.e. there is a single walker at distance d_{start}. The probability to find the random walker at distance d when starting at distance d_{start} after t time steps is then

$$p_d^{(d_{\text{start}})}(t) = \left(\boldsymbol{W}^t\right)_{d,d_{\text{start}}} \tag{5.3}$$

as the process is Markovian. This time evolution is displayed in Fig. 5.1 for $B = 8$ and $L = 3$. For a strong hierarchically clustered network ($\xi = 1/8$) one clearly sees the influence of the hierarchy layers when starting near the sink at $d_{\text{start}} = 1$. The random walker tends to be either absorbed, stay near ($d = 1$) or jump to a node at hierarchical distance $d = 2$. With progressing time there are phases where the walker tends to stay at certain distances and then leave those again for shorter distances. With growing structure parameter ξ, this behavior is smoothened until being at any distance is equally probable[2] for the Erdős–Rényi limit at $\xi = 8$. When starting 'far away' from the sink at distance $d_{\text{start}} = 3$ for the hierarchically clustered case ($\xi = 1/8$), the walker tends to stay far away, just slowly arriving at modules "nearer" to the sink until being absorbed.

[2] when corrected for the number of target nodes at distance d

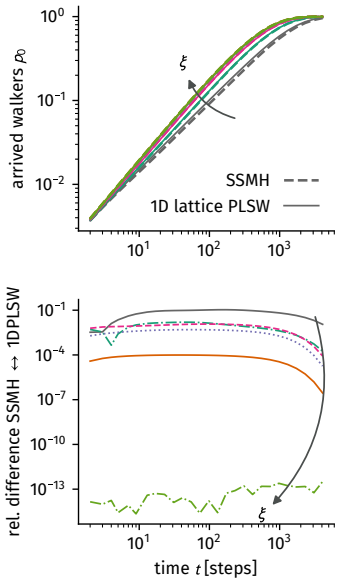


Figure 5.2: **(Top)** The SSMH average medium arrival time cdf $p_0(t)$ when starting anywhere in the network as per Eq. (5.4). Parameter values $B = 8$, $L = 3$, $\langle k \rangle = 7$ and $\xi \in \{0.125, 0.25, 0.5, 1, 4, 8\}$ were used. The results for the 1D lattice PLSW network model correspond to the SSMH results with increasing ξ as can be seen in the **(bottom)** panel.

Global Mean First Passage Time

We want to find the global mean first passage time (GMFPT) which quantifies the mean arrival time at sink node v for every node u as a start node. Suppose that a random walk begins with a single walker on every node $u \neq v$. The sink node v is chosen as the focal node and therefore positioned at hierarchical distance $d = 0$. Similar as to the procedure described in Sec. 3.1.4, one drops the row and column of transition matrix Eq. (5.1) which correspond to $d = 0$ to obtain a reduced transition matrix $\tilde{\boldsymbol{W}}$. Following the derivation presented in App. B.3.2, the ratio of walkers absorbed by the sink node v up to time t is given as

$$p_0(t) = 1 - \mathbf{1}^T \tilde{\boldsymbol{W}}^t \tilde{\boldsymbol{b}}. \tag{5.4}$$

Here, the vector $\mathbf{1}^T = (1, 1, ..., 1)$ was introduced, as well as the vector $\tilde{\boldsymbol{b}}$, which contains the fractions of all possible sources nodes at layer d as

$$\tilde{b}_d = \frac{B-1}{B^L - 1} B^{d-1}, \qquad 1 \leq d \leq L.$$

This arrival time cdf Eq. (5.4) is shown in Fig. 5.2 and compared to the cdfs of the corresponding 1D lattice PLSW model, which was computed using the full transition matrix composed from the 1D lattice PLSW pmf Eq. (4.25) with entries $W_{vu} = P_{d(u,v)} / \langle k \rangle$. The average media of both models produce qualitiatively similar arrival time cdfs which, with increasing structural control parameter ξ (i.e. increasing long-range connection probability), become even quantitatively similar.

After discussing the arrival time distribution, the focus is shifted to finding the global mean FPT for the focal node in the following. This quantity can be calculated as

$$\begin{aligned} \tau &= \sum_{t=0}^{\infty} t \left[p_0(t) - p_0(t-1) \right] = \sum_{t=0}^{\infty} t \left[\mathbf{1}^T \tilde{\boldsymbol{W}}^{t-1} \tilde{\boldsymbol{b}} - \mathbf{1}^T \tilde{\boldsymbol{W}}^t \tilde{\boldsymbol{b}} \right] \\ &= \mathbf{1}^T \left[\mathbb{1} - \tilde{\boldsymbol{W}} \right]^{-1} \tilde{\boldsymbol{b}}. \end{aligned} \tag{5.5}$$

Note that this quantity is equal to, both, the global mean FPT and the pair-averaged FPT (that is, the mean global mean FPT), because in this AMA all nodes are equal. This result is similar to the result for arbitrary networks, where instead of $(\mathbb{1} - \tilde{\boldsymbol{W}})^{-1}$ one makes use of the inverse of the reduced unnormalized graph Laplacian [32, pp. 147 ff.]. However, using the layer approach, the matrix size can be reduced from $B^L - 1$ to L, a great reduction in degrees of freedom. This makes it possible to obtain analytical expressions for the global mean FPT. As an example, the result for $L = 3$ is

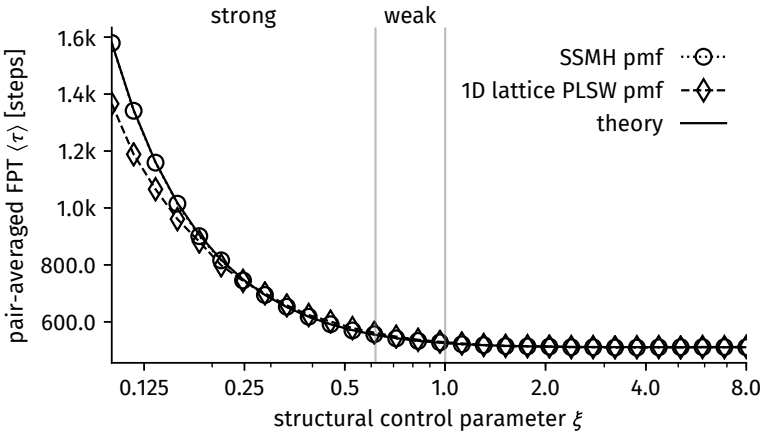


Figure 5.3: Pair-averaged FPT for the SSMH model and the 1D lattice model computed from simulations (markers) and compared to Eq. (5.6). Parameters used here where $B = 8$, $L = 3$ and $\langle k \rangle = 7$. The figure was reproduced from [93].

$$\begin{aligned}\tau_{L=3} = \Big[&(B-1)B^2\left(\xi^2+\xi+1\right)\times\\ &\times\Big(B^4\xi^2(\xi+1)+B^3\xi(\xi+1)+\\ &\quad+B^2(\xi+1)^2-B\xi\left(2\xi^2+3\xi+2\right)+\xi^2(\xi+1)\Big)\Big]\\ \Big/\Big[&\left(B^2+B+1\right)\xi^2(B\xi+B-\xi)\times\\ &\times\left(B\left(\xi^2+\xi+1\right)-\xi(\xi+1)\right)\Big].\end{aligned} \tag{5.6}$$

Note that for $\xi \to B$ the global mean FPT approaches the avergage medium Erdős–Rényi solution $\tau = B^L - 1$, which corresponds to the complete network. In addition, the global mean FPT diverges for $\xi \to 0$, since the average medium then approaches a state where it consists of B^{L-1} complete networks, each containing B nodes.

The arrival time pmf is shown in Fig. 5.4 and compared to the pmfs of the corresponding one-dimensional (1D) lattice PLSW model, which was computed using the full transition matrix composed from the 1D lattice PLSW pmf Eq. (4.25) with entries $W_{vu} = P_{d(u,v)}/\langle k \rangle$, as well as the approximative relationship $p_0(t) - p_0(t-1) = \tau^{-1}\exp(-t/\tau)$ where τ is computed from Eq. (5.6). The average media of both models produce qualitatively similar arrival time cdfs which, with increasing structural control parameter ξ (i.e. increasing long-range connection probability), become even quantitatively similar. The computed pmf for the SSMH model corresponds to the approximative relationship, alas showing systematic deviations.

A comparison of simulated pair-averaged FPT and Eq. (5.6) can be seen in Fig. 5.3. Here, the pair-averaged FPT was computed for average media of SSMH and the 1D lattice PLSW model using parameters $B = 8$, $L = 3$, $\langle k \rangle = 7$ and structural control parameter $0.1 \leq \xi \leq 8$. Furthermore, random walk simulations were performed using averaged networks where at initial time $t = 0$ a single walker was placed on every node. The simulation was then run as described in Sec. 3.1.1 until each walker visited each node at least once. At every time step t a walker visited a node it had not visited before, this

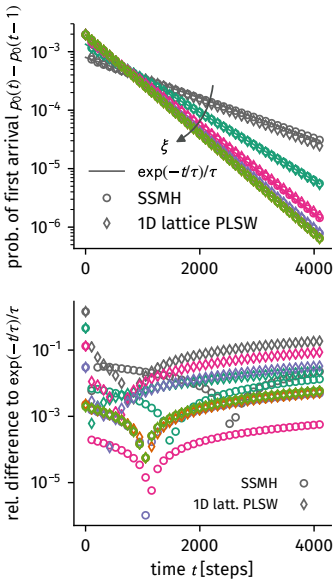


Figure 5.4: **(Top)** the SSMH average medium arrival time pmf$p_0(t) - p_0(t-1)$ as per Eq. (5.4) as well as the one-dimensional lattice PLSW average medium arrival time pmf computed from its full transition matrix. Parameter values $B = 8$, $L = 3$, $\langle k \rangle = 7$ and $\xi \in \{0.125, 0.25, 0.5, 1, 4, 8\}$ were used. The pmfs approximately follow an exponential distribution with mean Eq. (5.6). **(Bottom)** relative difference of the pmf of both models to the approximative relationship.

time t was saved as a FPT. The pair-averaged FPT was computed as the mean of all saved FPTs. The simulations are in agreement with the theory. For small values of $\xi \lesssim 0.2$ the results for SSMH and PLSW begin to differ more strongly. Note that the pair-averaged first passage time is a monotonically decreasing function of the structural control parameter ξ which suggests that no 'optimal' structure exists in between the two limiting structures—the global minimum lies in the randomized limit.

5.1.2 *Mixing Time and Small-World Effect in the Modified Watts–Strogatz Model*

In this section, an average medium approximation will be used to provide an intuitive and analytic illustration of the original small-world effect as introduced in [1]. Parts of this subsection are close to the text first published in [94].

In the original model, the small-world effect was established by comparing the clustering coefficient to the average shortest path length of the networks. While random networks have short path lengths, they possess low clustering, on the other hand regular networks are highly clustered, while nodes are, on average, quite distant from one another. With rewiring only a short amount of edges it was shown that shorter paths appear immediately while high clustering is preserved, explaining the small-world effect. However, in the following, a different approach will be taken.

One of the purposes of the original model was to explain the Milgram small-world experiment [8] where participants had to mail packets[3] to a stranger by mailing them to a person they did know on a first-name basis and instruct them to pass the packet further. Even though this experiment had methodological flaws [121, 122], it found that some of the packets arrived relatively fast through a short chain of people. In the following the small-world effect will be illustrated by showing that the mixing time, as an upper bound for the delivery time of those messages, decreases much faster than the clustering coefficient with increasing probability of long-range edges.

[3] also referred to as *messages* in the following

As argued above, the average medium random walk transition matrix is given as $W_{vu}^{\text{avg}} = \langle A_{vu} \rangle / k = P_{d(v,u)}/k$ where the modified small-world connection probability Eq. (4.12) is used. The time scale with which the equilibrium distribution is approached on this average medium network is given by the eigenvalue gap of the transition matrix W_{vu}^{avg} as $t_{\text{mix}}^{-1} = 1 - \omega_1$ where $\omega_0 = 1$ is the largest eigenvalue and ω_1 is the second largest eigenvalue as introduced in Sec. 3.1.3. The average medium transition matrix is circulant based on the vector

$$\boldsymbol{w} = k^{-1}(0, \underbrace{p_S, \ldots, p_S}_{k/2}, \underbrace{p_L, \ldots, p_L}_{N-1-k}, \underbrace{p_S, \ldots, p_S}_{k/2}).$$

In this case, the j-th eigenvalue of W_{vu}^{avg} is given as

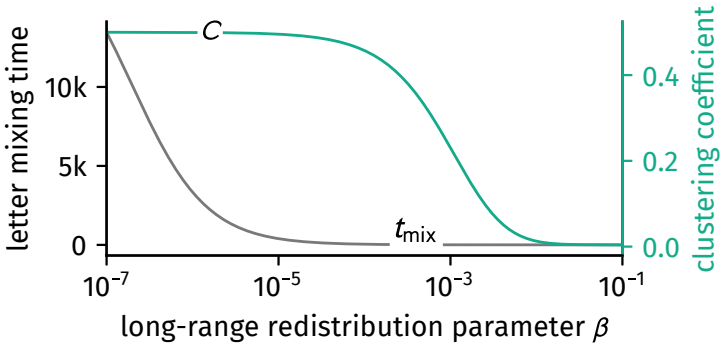


Figure 5.5: The small-world effect as illustrated by the modified Watts-Strogatz clustering coefficient and mixing time. The random walk message delivering mixing time Eq. (5.7) as an upper bound of the real message delivering mixing time decreases rapidly with the introduction of long-range links while clustering Eq. (4.17) preserves. Displayed here are results for $N = 1001$ and $k = 4$. The figure was taken from [94].

$$\omega_j = \sum_{v=1}^{N} w_v \exp(i2\pi v/N)$$

such that the second largest eigenvalue can be computed as $\omega_1 = p_S\Gamma/k - p_L(1+\Gamma)/k$ where

$$\begin{aligned}\Gamma &= 2\sum_{j=1}^{k/2}\cos(2\pi j/N)\\ &= k - \left(\frac{\pi}{N}\right)^2\frac{k}{3}(k/2+1)(k+1) + \mathcal{O}(N^{-4})\end{aligned}$$

which yields the mixing time

$$t_{\text{mix}}(\beta) = \left[1 - \frac{\Gamma - \beta(1+\Gamma)}{k+\beta(N-1-k)}\right]^{-1}. \tag{5.7}$$

Fig. 5.5 displays how both clustering coefficient and mixing time decrease with increasing long-range redistribution parameter β. In the limits one finds the results from k-regular networks and the average medium approximation of the Erdős–Rényi graph

$$\begin{aligned}t_{\text{mix}}(\beta=0) &= \left[1-\frac{\Gamma}{k}\right]^{-1} \overset{N\gg k/2}{\longrightarrow} \frac{N^2}{\pi^2}\frac{3}{(k/2+1)(k+1)},\\ t_{\text{mix}}(\beta=1) &= \left[1-\frac{1}{N-1}\right]^{-1} = 1-\frac{1}{N}.\end{aligned}$$

This implies that for small long-range redistributions the relative mixing time decreases as

$$\frac{t_{\text{mix}}(\beta)}{t_{\text{mix}}(0)} = 1 - \beta\left(\frac{3N^3}{\pi^2\,(k/2+1)\,(k+1)k} - \frac{N}{k} + \frac{1}{k}\right) + \mathcal{O}(\beta^2). \tag{5.8}$$

Comparing Eqs. (4.18) and (5.8), one can see that for small β the rate with which the mixing time decreases is of order N^3 while the rate with which the clustering coefficient decreases is of order N, which is a difference of two orders of magnitude. This shows that even with a small amount of long-range connection probability, the delivery time of randomly passed messages declines rapidly while clustering is still preserved. Note that the mixing time does not possess a minimum when varying the model's structural control parameter.

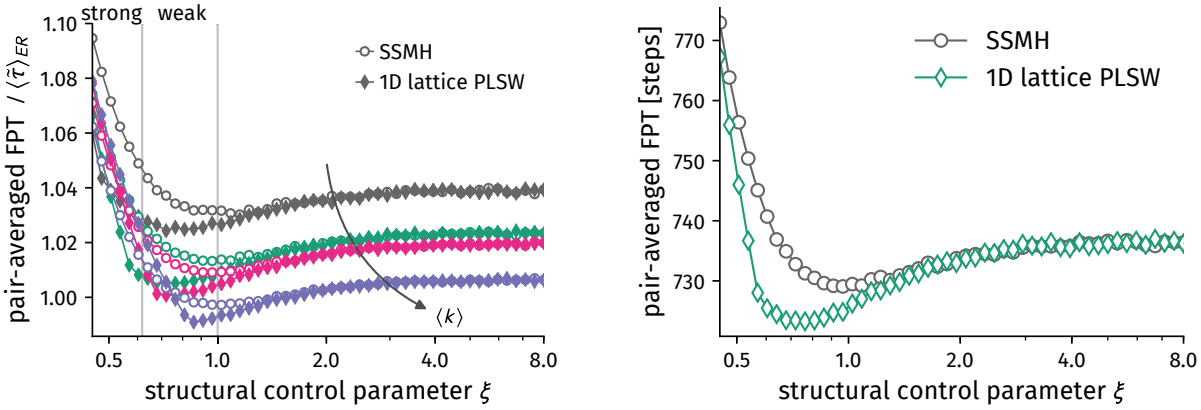


Figure 5.6: Pair-averaged first passage time computed numerically using Eq. (3.11) for the self-similar modular hierarchical (SSMH) and the 1D lattice PLSW network models introduced in the previous chapter. **(Left)** Pair-averaged FPT as a function of the structural control parameter ξ for $B = 8$, $L = 3$, and $\langle k \rangle \in \{6, 7, 8, 10\}$ (top to bottom curves, respectively), averaged over 2000 SSMH and 1D lattice PLSW network realizations and normalized by the lower bound of the pair-averaged FPT on Erdős-Rényi networks given in Eq. (5.11). **(Right)** Unnormalized result for $\langle k \rangle = 7$. The figure was reproduced from [93].

5.2 *Actual Network Realizations*

So far, theoretical results have been derived for average media, finding that neither mixing time nor pair-averaged FPT seem to be optimal in between the clustered limit and the random limit. In nature, media are not averaged. Instead, one finds structures where connections exist or do not exist. Any information spreading on these networks has to follow discrete paths between nodes to reach other nodes to which its originating node is not connected. In this subsection, network samples from the models introduced in the previous chapter are investigated to see whether the average medium approximation used before fails to describe effects provided by discrete structure.

It will be found that those models from Ch. 4 which are based on lattice node positions indeed show a minimum in the pair-averaged FPT in between both limit structures, while the 2-torus PLSW model does not. A heuristic explanation of this effect is found by deriving a formula to estimate the pair-averaged FPT in a locally clustered environment, which is composed of the clustering coefficient and the degree variance.

This section is close to the text first published in [93].

5.2.1 *Numerical Results for Pair-Averaged First Passage Time*

Similarly to the procedures of the previous sections and the previous chapter, the parameters $B = 8$ and $L = 3$, where chosen, here with mean degree values of $\langle k \rangle \in \{5, 6, 7, 8, 9, 10, 11\}$. Subsequently, different values of the structural control parameter $0.5 \lesssim \xi \leq B$ (or the corresponding μ parameter) were scanned. Note that the lower bound $0.5 \lesssim \xi$ was chosen here because at this point the pair-averaged FPT is already much greater than its value from the corresponding Erdős–Rényi limit, thus no remarkable information can be gained from further decreasing ξ, other than that it will increase even more.

For each parameter set, 2000 different SSMH network realizations were generated with $N = B^L$ nodes, using the algorithm described in Sec. 4.1. Similarly, 2000 PLSW network realizations were sampled with $N = B^L$ nodes, the same mean degree $\langle k \rangle$ values, and the corresponding structural control parameter $\mu = \log \xi / \log B$, using the algorithm described in Sec. 4.3.3. After selecting the largest component of each of the resulting networks, the pair-averaged FPT $\langle \tau \rangle$ was measured using the eigenvalues and eigenvectors of the unnormalized graph Laplacian, as described in Sec. 3.1.5.

As one may observe in Fig. 5.6, a minimum in the pair-averaged FPT emerges in the $\xi \lesssim 1$ region for both the SSMH and the 1D lattice PLSW networks. This corresponds to the point where the network structures leave the weak hierarchical clustering regime in which an average node has more connections to nodes in lower hierarchical layers than to nodes in higher layers, but where both quantities are of similar order.

Furthermore, the modified Watts-Strogatz model was tested to display a minimum in the pair-averaged FPT. The parameters $B = 8$, $L = 3$, $\langle k \rangle = 10$ where chosen and the structural control parameter $0.5 \lesssim \xi \leq 8$ was varied. The pair-averaged FPT for the largest component of each network sample was found numerically using Eq. (3.11) and finally averaged over 10^4 independent samples each.

Finally, the 2-torus PLSW model, while structurally interpretable to be modular hierarchical, does not display a minimum in the pair-averaged FPT. Eq. (3.11) was used on the largest component of 10^4 independent samples each for $N = 8^3$ and varying mean degree $\langle k \rangle \in \{6, 8, 12\}$, as well as increasing structural control parameter $-\kappa$. The result can be seen in Fig. 5.7.

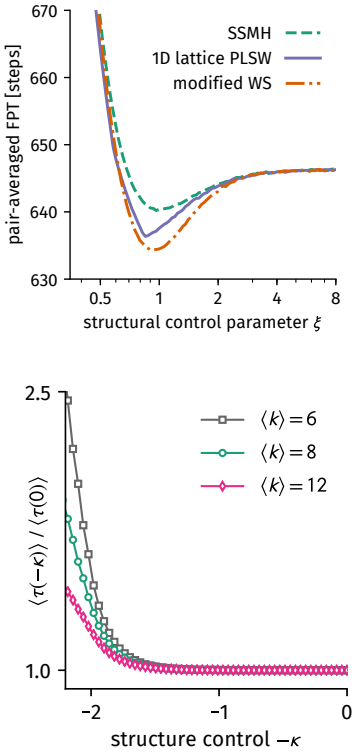


Figure 5.7: Pair-averaged FPT for network samples of the main small-world models introduced in Ch. 4. **(Top)** The modified Watts-Strogatz, self-similar modular hierarchical as well as the 1D lattice power-law small-world model all display a minimum for an intermediate degree of long-range connection probability. **(Bottom)** The 2-torus power-law small-world model does not show a minimum in the pair-averaged FPT.

5.2.2 *Heuristic Explanation of Minimum*

Given that the AMA approach described above fails to reproduce the minimal diffusion time observed numerically at intermediate levels of hierarchical modularity (see Fig. 5.6), a heuristic approach is developed below that will help understand the origin of this effect.

The following calculation is based on the heuristic lower bound of the global mean FPT of a target node v as introduced in Sec. 3.1.5. The base assumption is that the network is locally tree-like (i.e., that it has vanishing clustering coefficient) and that the mean FPT for any target node v is exponentially distributed as $\exp(-t\beta_v)$, with

$$\beta_v = \frac{k_v}{N \langle k \rangle} \left(1 - \frac{1}{\langle k \rangle}\right).$$

Since the mean FPT is only asymptotically exponentially distributed, by using this assumption only a lower bound will be obtained for the corresponding pair-averaged FPT. One thus finds $\langle \tau \rangle \geq N^{-1} \sum_i \beta_i^{-1}$, which yields

$$\langle \tau \rangle \geq N \langle k \rangle \left\langle \frac{1}{k} \right\rangle_{k>0} \frac{1}{1 - \langle k \rangle^{-1}}, \tag{5.9}$$

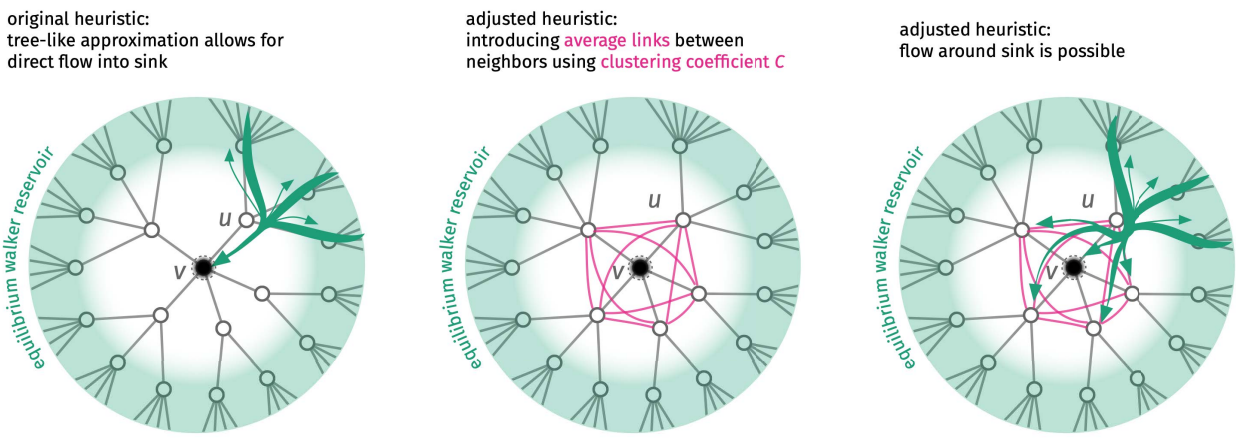


Figure 5.8: Illustration of the heuristic method to find the global mean first passage time (FPT) of a sink node v with non-zero clustering. For both the original and the adjusted heuristic, all nodes which are more than two steps away from the sink node are considered to carry an amount of remaining walkers which is proportional to their equilibrium walker density. **(Left)** In the original heuristic, a locally tree-like network is assumed and hence, walkers can flow from the reservoir into first neighbors, displayed here for an example node u. From there they can either flow into the sink or back into the reservoir. **(Center)** In the adjusted heuristic, local clustering is considered by introducing average links between all neighbors of v using the clustering coefficient C. **(Right)** Due to the added links between first neighbors, walkers can flow around the sink before being absorbed or flowing back into the reservoir, hence increasing the arrival time.

where $\langle\cdot\rangle_{k>0}$ denotes the average over all nodes with non-zero degree. Even though tree-like approximations are often remarkably useful even in clustered networks, they lose their validity when the average shortest path length becomes too distinct from the average path length in a random network [102], which is the case for the network models from Ch. 4. Hence, an extension of this result is derived in the following, which explicitly considers a non-vanishing clustering coefficient.

The approach taken in [100] treats the time-dependent walker concentration $p_u(t)$ on any node u at time t as composed of two separable distributions

$$p_u(t) \approx p_u p_{\text{total}}(t).$$

and finds that the time-independent walker concentration on neighboring nodes u of sink node v follows as

$$p_u = \sum_{u' \in \text{Nei}(u)\setminus v} \frac{p_{u'}}{k_{u'}} = \frac{1}{N\langle k\rangle}(k_u - 1)$$

when the networks are locally tree-like. A better approximation for locally clustered networks is given as follows, where the argumentation follows Sec. 3.1.5 and picks up at Eq. (3.13). It is furthermore illustrated in Fig. 5.8.

Consider sink node v with degree k_v. On average, a neighbor u of v will have $c_v = C_v(k_v - 1)$ neighbors that are also neighbors of v. The rest of its $k_u - 1 - c_v$ neighbors each contribute an influx of $(N\langle k\rangle)^{-1}$ walkers, such that

$$p_u = \frac{1}{N\langle k\rangle}\left[k_u - 1 - c_v\right] + \sum_{u' \in \text{Nei}(u)\setminus v} \frac{p_{u'}}{k_{u'}}. \tag{5.10}$$

Now, every neighbor u is replaced with an average node of degree $\langle k\rangle$. In addition, every node i is assumed to have the same number of edges between neighbors, so $c_v \equiv c = C(\langle k\rangle - 1)$. This transforms Eq. (5.10) into a self-consistent expression given by

$$p_u = \frac{1}{N\langle k\rangle}(\langle k\rangle - 1 - c) + c\frac{p_u}{\langle k\rangle},$$

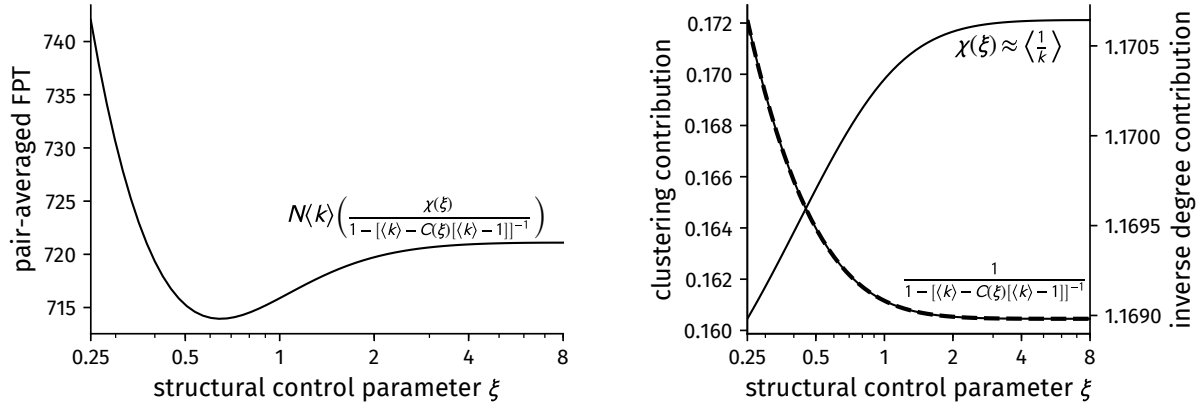


Figure 5.9: Heuristic result **(a)** for the pair-averaged FPT as a function of the structural control parameter ξ, as given by Eq. (5.11) with parameters $B = 8$, $L = 3$ and $\langle k\rangle = 7$, and **(b)** the mean inverse degree and clustering coefficient contributions to this curve, showing how the combination of these growing and decaying functions can result in the observed minimum shown in Fig. 5.6. The figure was reproduced from [93].

which is equivalent to

$$p_u = \frac{1}{N}\left(1 - \frac{1}{\langle k\rangle - C\big[\langle k\rangle - 1\big]}\right).$$

Using Eq. (3.12), the decay rate is then found to be

$$\beta_v = \frac{k_v}{N\langle k\rangle}\left(1 - \frac{1}{\langle k\rangle - C\big[\langle k\rangle - 1\big]}\right).$$

One therefore finds that the global mean FPT of target node v is approximately lower bounded by

$$\tau_v \geq \frac{N\langle k\rangle}{k_v}\left(\frac{1}{1-\Big[\langle k\rangle - C\big[\langle k\rangle - 1\big]\Big]^{-1}}\right). \tag{5.11}$$

The pair-averaged FPT (i.e. the mean global mean FPT) $\langle\tau\rangle = (1/N)\sum_v \tau_v$ is consequently lower bounded by

$$\langle\tau\rangle \geq N\langle k\rangle \left\langle\frac{1}{k}\right\rangle_{k>0}\left(\frac{1}{1-\Big[\langle k\rangle - C\big[\langle k\rangle - 1\big]\Big]^{-1}}\right). \tag{5.12}$$

In order to quantitatively evaluate this result for the SSMH model one needs to compute the mean inverse degree $\langle k^{-1}\rangle_{k>0}$, which can be numerically intensive. A further approximation of the result is given by expressing this quantity by its second order Taylor expansion around the mean $\langle k\rangle$, such that

$$\begin{aligned}\left\langle\frac{1}{k}\right\rangle_{k>0} \approx \chi(\xi) &= \frac{1}{\langle k\rangle} + \frac{\langle k^2\rangle}{\langle k\rangle^3} + \Phi_{ER} \\ &= \frac{\langle k^2\rangle - \langle k^2\rangle_{ER}}{\langle k\rangle^3} + \left\langle\frac{1}{k}\right\rangle_{ER,k>0}.\end{aligned} \tag{5.13}$$

For the result Eq. (5.12) the expressions for the clustering coefficient C in Eq. (4.11) and for the second moment of the degree distribution in Eq. (4.7) are used. The constant Φ_{ER} contains the error made

by considering networks without nodes of degree $k = 0$, which alters the degree distribution in comparison to the second moment $\langle k^2 \rangle$. This issue can be fixed by demanding $\chi(\xi = B) = \langle k^{-1} \rangle_{ER,k>0}$, which can be found by averaging k^{-1} for $k > 0$ using a binomial pmf with $n = B^L$ and $p = k/(B^L - 1)$, yielding

$$\Phi_{ER} = \left\langle \frac{1}{k} \right\rangle_{ER,k>0} - \frac{1}{\langle k \rangle} - \frac{\langle k^2 \rangle_{ER}}{\langle k \rangle^3}.$$

The function Eq. (5.12) is plotted in Fig. 5.9a for the SSMH model. Its shape as a function of ξ results from two opposing effects (see Fig. 5.9b): the growth of the mean inverse degree and the decay of the clustering coefficient (as derived in Sec. 4.1.2) for increasing ξ. The mean inverse degree as approximated by $\chi(\xi) \propto \langle k^2 \rangle$, grows monotonically with ξ, as shown in Sec. 4.1.2. This results from the fact that the shape of the degree distribution is rather narrow for small ξ values, which leads to a vanishing number of nodes with small degrees. At the same time, this monotonic growth is countered by the clustering behavior because, as the structural control parameter ξ is increased, the network loses its strong clustering and thus the clustering contribution decreases. The combination of these two curves leads to the appearance of a minimum in the pair-averaged FPT at intermediate levels of hierarchical modularity, as shown in Fig. 5.9a. The reader should note, however, that the position of the minima given by Eq. (5.11) differs from those in Fig. 5.6, due to the heuristic origin of this equation.

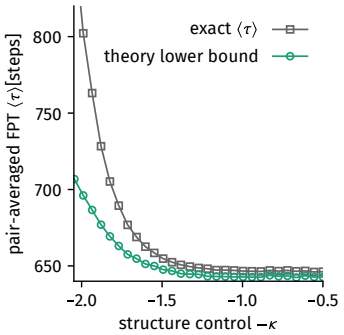


Figure 5.10: Comparison between the actural pair-averaged FPT and the heuristic lower bound on a 2-torus PLSW model with $N = 8^3$ and $\langle k \rangle = 10$ The exact value was obtained using Eq. (3.11) and the lower bound was computed using Eq. (5.12). Markers represent the averages over 1000 independent network samples for each of which the largest component was extracted. The lower bound uses the mean inverse degree $\langle k^{-1} \rangle$ and the mean clustering coefficient C which were computed from those largest components, as well.

The structure of Eq. (5.11) further explains why the networks generated from the 2-torus PLSW model do not show a minimum in the pair-averaged FPT was shown in Fig. 5.7. As argued above, the minimum observed for the other models emerges from a varying degree variance which increases with increasing amount of long-range connection probability. However, it was shown in Sec. 4.3 that the 2-torus PLSW model has constant degree variance over the total range of its structural control parameter (see Fig. 4.10). Hence the minimum cannot exist as only the clustering contribution of Eq. (5.11) plays a role. A comparison of the exact pair-averaged FPT Eq. (3.11) and the lower bound Eq. (5.12) is shown in Fig. 5.10.

One might wonder how necessary the extension of the heuristic to non-zero clustering actually is. Considering the original result Eq. (5.9) it now becomes clear that by considering locally tree-like networks, only the inverse degree contribution would influence the shape of the pair-averaged FPT as a function of the structural control parameters and thus it would remain constant for the 2-torus PLSW. For the remaining models it would *increase* with increasing structural control parameters because the inverse mean degree increases, directly contradicting the numerical results.

It should finally be noted that even though the adjusted heuristic produces a result which is qualitatively in agreement with the numerically evaluated pair-averaged FPT, the difference between the heuristic lower bound and the exact value increases for all models

when approaching the regular clustered limit (i.e. for low values of the structural control parameter where the networks become increasingly lattice-like).

5.3 Summary and Discussion

Within this chapter the hypothesis of modular hierarchical networks being optimal for a fast pair-averaged first passage time in random walks was tested. At first, an average medium approximation was used to find that average modular hierarchical and other small-world models do not possess minimal passage or mixing times. It was shown that the SSMH and one-dimensional (1D) lattice PLSW model closely correspond considering their passage-time statistics, reflecting the result from Ch. 4 that they are of similar structure.

Subsequently it was shown that when analyzing actual networks sampled from the models, the pair-averaged first passage time does indeed show a minimum for the SSMH, 1D lattice PLSW and modified Watts-Strogatz model, an effect not replicated by the average medium. In order to explain the minimum, a heuristic formula for the pair-averaged first passage time on networks with non-zero local clustering was derived. Following this heuristic, the minimum emerges as a combined effect of increasing degree variance and decreasing clustering coefficient with increasing long-range connection probability. This highlights the fact that average medium approximations may yield contradicting results when considering spreading processes on actual networks.

The increasing node degree variance, however, seems to be an artifact stemming from discrete node positions in the SSMH, 1D lattice PLSW and modified WS model. In the 2-torus PLSW model, which is equally capable of producing networks describable as hierarchically modular, the node degree variance is constant. Consequently the minimum vanishes for this network model.

Hence, an optimal structure for fast diffusion can be expected of networks which are in a state where both their degree variance and clustering coefficient are low compared to other structures, resulting in a minimum pair-averaged FPT. However, no such statement can be made in general for modular hierarchical networks, in contrast to the results of informed searches (see Sec. 2.2.4).

Concerning the discrepancy between the lower bound pair-averaged FPT and the numerically evaluated pair-averaged FPT, this discrepancy is expected to arise since the adjusted heuristic still makes the assumption that the random walk process equilibrates in a few number of time steps $t_{\text{mix}} \ll N$. In networks which are structurally close to lattices, the shortest path length between pairs of nodes is substantially larger compared to more random networks and so the equilibration process will take much longer. One might expect that in a further adjusted heuristic, the increasing path length of lattice-like networks will play a role. This is connected to the observation that tree-like approximations of networks work reasonably well even

in clustered networks if the shortest path length is small [102]. One might therefore expect the shortest path length to play an even larger role than the clustering coefficient.

6
Cover Time for Arbitrary Complex Networks

For a discrete-time random walk, the cover time refers to the number of steps a random walker takes to visit each node in a network at least once (see Sec. 3.1.6). It places an upper bound on exhaustive search processes and the time to distribute a commodity to every node in a network. Since neither a simple method nor a heuristic exist for reliably computing this observable, a heuristic method is introduced in the following. In Sec. 6.1, the heuristic approach is explained and used to derive the cover time probability distribution as well as its first moment. Subsequently, the heuristic is applied to compute the mean cover time for various network models and real-world networks in Sec. 6.2. The analysis is finally concluded with a detailed error assessment in Sec. 6.3, alongside which it will be revealed that some of the small-world network models introduced in Ch. 4 possess a minimum cover time in the strong hierarchically clustered regime while the 2-torus PLSW does not. Both will be explained using the heuristic introduced in Sec. 5.2.2. Secs. 6.1–6.3 are close to the text first published in [92],

6.1 Heuristic Evaluation Method

6.1.1 Derivation

As has been argued in Sec. 3.1.5, given that a random walk process on a network equilibrates quickly, the information about the initial walker position becomes irrelevant rather rapidly. Here, fast equilibration refers to the fact that the initial random walker density on nodes of the network approaches the equilibrium density Eq. (3.3) in a small number of time steps $t_{\mathrm{mix}} \ll N$. As shown in [100] and argued in Sec. 3.1.5, the first passage time at a target node v is then distributed asymptotically according to

$$p_v(\tau) \propto \exp(-\tau/\tau_v) \tag{6.1}$$

where τ_v is the global mean FPT of Eq. (3.10). In this case, the asymptotic behavior refers to larger FPTs τ. The global mean FPT τ_v can

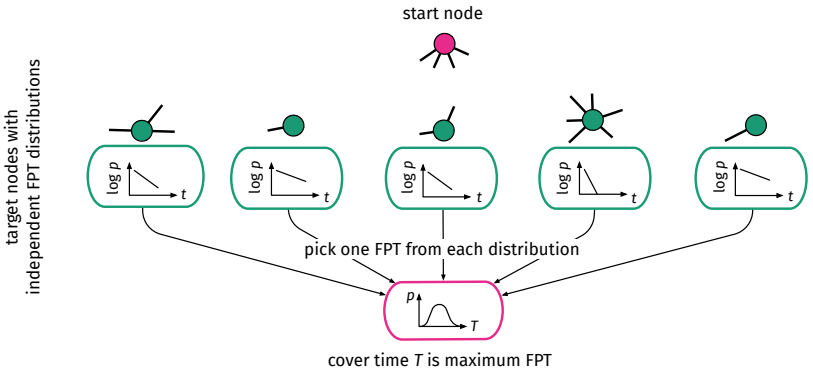


Figure 6.1: Illustration of the heuristic approach. Each node of the network is represented as an independent object that can be visited by the walker starting at the pink node. Each target node v is associated with its own first passage time (FPT) distribution. This FPT distribution is asymptotically distributed as $\propto \exp(-\tau/\tau_v)$ (for larger times τ) with its global mean FPT τ_v. In order to compute the cover time, one FPT τ is drawn from every target's distribution. The cover time is then given as the maximum time of all drawn FPTs. The figure was reproduced from [92], © 2017 American Physical Society.

vary for each node and only depends on the structural properties of the network.[1]

[1] Note that in the following derivations, the FPT decay rate $\beta_v = \tau_v^{-1}$ will be often used instead of the global mean FPT to simplify the notation.

Following the heuristic approach presented below, one can proceed as illustrated in Fig. 6.1 to obtain, first, the cover time distribution and second, the mean cover time, both based on the collection of target FPT distributions $p_v(\tau)$. Ignoring the start node u at which a walker begins its random walk, an FPT t_v is drawn for each potential destination v from their respective target FPT distribution p_v, resulting in the set $\mathcal{F}$ which contains all $N-1$ target FPTs. The cover time T_u is then given as the maximum element of the set $\mathcal{F}$. The probability that a time T is an upper bound of $\mathcal{F}$ can be found as the probability that no element of $\mathcal{F}$ is greater than T, which is

$$\begin{aligned} P_u(T) &= P_u(\text{“}T \text{ is an upper bound of all target FPTs”}) \\ &= P_u(t_v \leq T \ \forall t_v \in \mathcal{F}) = \prod_{v \neq u} P(t_v \leq T) \end{aligned} \tag{6.2}$$

and is equal to the cumulative distribution function (cdf) of T. Assuming the validity of Eq. (6.1), the discrete time probability mass function for first passage time τ at node v is evaluated as

$$\begin{aligned} p_{v,\tau} &= \exp(-\beta_v \tau) \Big/ \sum_{t=1}^{\infty} \exp(-\beta_v t) \\ &= [1 - \exp(-\beta_v)] \exp(-\beta_v(\tau - 1)), \end{aligned}$$

yielding the cdf

$$\begin{aligned} P(t_v \leq T) &= \sum_{t=1}^{T} p_{v,t} = (1 - \exp(-\beta_v)) \sum_{t=1}^{T} \exp(-\beta_v(t-1)) \\ &= 1 - \exp(-\beta_v T). \end{aligned} \tag{6.3}$$

Eqs. (6.2) and (6.3) yield the cumulative distribution function for the cover time,

$$P_u(T) = \prod_{v \neq u} \Big(1 - \exp(-\beta_v T)\Big).$$

The cdf can be further approximated by assuming a continuous time distribution, which simplifies the following derivations while the outcome does not change significantly, as explained in Sec. 6.3, such

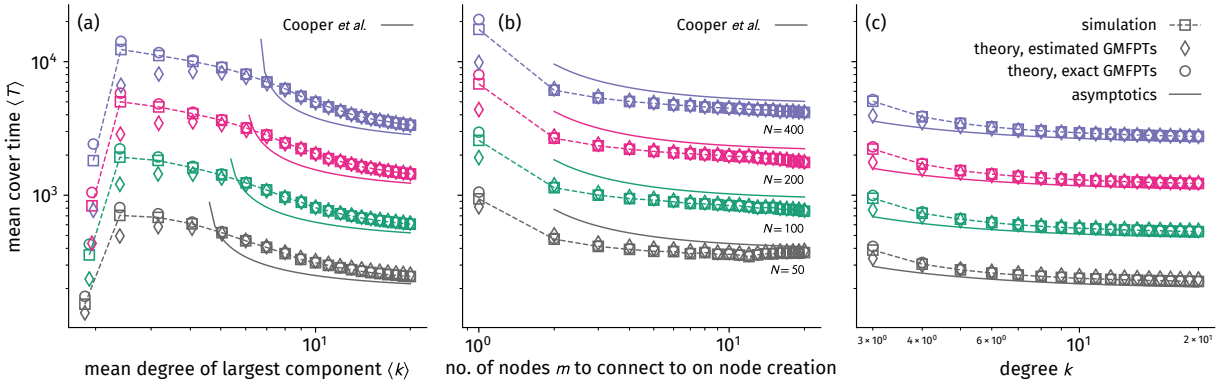


Figure 6.2: The mean cover time $\langle T \rangle$ of the largest component of **(a)** Erdős-Rényi (ER) networks, **(b)** Barabási-Albert (BA) networks , and **(c)** random k-regular networks. Shown are averages over 1000 simulations per data point (□) and both estimations of the mean cover time using (♢) estimated GMFPTs from the target nodes' degrees Eq. (6.9) and (◯) exact GMFPTs Eq. (6.9) computed from the unnormalized graph Laplacian's spectrum in (a) and (b). Respectively, Eq. (6.10) has been used for random k-regular networks in (c). Dashed lines are simulation results and a guide to the eye. Further displayed are the asymptotic results derived for the ER [111] and BA [112] network models as well as the lower bound Eq. (6.11). The figure was reproduced from [92], © 2017 American Physical Society.

that the mean cover time of start node u is approximately

$$T_u \approx \int_0^\infty \mathrm{d}T \left[1 - \prod_{v \neq u} \left(1 - \exp(-\beta_v T)\right)\right] - \frac{1}{2}. \tag{6.4}$$

Note that a bias of $1/2$ emerges when changing from discrete time to continuous time, as discussed in Sec. 6.3. However, the mean cover time is lower bounded by the cover time of a complete graph which scales as $N \log N$, such that one usually finds $T_u \gg 1/2$. Hence, the bias will be omitted in the following, which introduces relative error of size $(2T_u)^{-1}$. Now, the global mean cover time can be found by taking the average over all target nodes u as

$$\langle T \rangle \approx \int_0^\infty \mathrm{d}T \left[1 - P(T) \frac{1}{N} \sum_{u=1}^{N} \frac{1}{1 - \exp(-\beta_u T)}\right] \tag{6.5}$$

where

$$P(T) = \prod_{v=1}^{N} \left(1 - \exp(-\beta_v T)\right). \tag{6.6}$$

As shown in Sec. 6.3.3, introducing small relative error of order $\mathcal{O}\left((N \log N)^{-1}\right)$ for the networks discussed in this paper, one can further simplify this integral to find

$$\begin{aligned} \langle T \rangle &\approx \int_0^\infty \mathrm{d}T \left[1 - P(T)\right] = \int_0^\infty \mathrm{d}T \left[1 - \prod_{v=1}^{N} \left(1 - \exp(-\beta_v T)\right)\right] \\ &= \sum_{\mathcal{S} \in \mathcal{P}^*(\mathcal{V})} (-1)^{|\mathcal{S}|+1} \left(\sum_{v \in \mathcal{S}} \beta_v\right)^{-1}, \end{aligned} \tag{6.7}$$

where $\mathcal{V}$ is the set containing all nodes and $\mathcal{P}^*(\mathcal{V})$ is the set of all possible subsets of $\mathcal{V}$ (excluding the empty set). Conceptually, this integral corresponds to a situation where an additional node is inserted on which every random walk starts but which can never be visited again. Even though the integral Eq. (6.7) can be solved analytically to obtain the result above, in practice it is more feasible to solve the integral numerically than iterating over $\mathcal{P}^*(\mathcal{V})$ which has $2^N - 1$ elements and hence becomes very large rather quickly.

Now, the estimation of the global mean cover time reduces to an efficient estimation of the FPT decay rates β_v. There are two ways to estimate the decay rates with the GMFPTs as described in Sec. 3.1.6. Using the estimation of the lower bound Eq. (3.14), the estimated global mean cover time is given by

$$\langle T^{\text{es}} \rangle \geq \int_0^\infty \mathrm{d}T \left[1 - \prod_{v=1}^{N} \left(1 - \exp\left(-T k_v \frac{1 - \langle k \rangle^{-1}}{N \langle k \rangle} \right) \right) \right]. \tag{6.8}$$

The advantage of this method is that only the network's degree sequence k_v needs to be known in order to estimate the global mean cover time. However, this method can obviously only account for a lower bound. One can also compute the exact GMFPTs using Eq. (3.11). In this case the computed global mean cover time is

$$\langle T^{\text{ex}} \rangle = \int_0^\infty \mathrm{d}T \left[1 - \prod_{v=1}^{N} \left(1 - \exp\left(-\frac{T}{\tau_v^{\text{ex}}} \right) \right) \right]. \tag{6.9}$$

6.1.2 *Cover Time of Networks With Equal GMFPTs*

Considering a network in which all nodes have approximately the same GMFPT $\tilde{\tau}$ and on which a random walk equilibrates quickly ($t_{\text{mix}} \ll N$), the mean cover time can be estimated using Eq. (6.4) to obtain

$$\begin{aligned} \langle T \rangle (\tilde{\tau}) &\approx \int_0^\infty \mathrm{d}T \ \left[1 - (1 - \exp(-T/\tilde{\tau}))^{N-1} \right] \\ &= \tilde{\tau} \Big[\gamma + \psi(N) \Big], \end{aligned} \tag{6.10}$$

where $\gamma \approx 0.57722$ is the Euler-Mascheroni constant, $\psi(z) = \Gamma'(z)/\Gamma(z)$ and $\Gamma(z)$ the gamma function.

An example for networks fulfilling the conditions above are random k-regular networks as introduced in Sec. 2.2.1 where all nodes have identical degree and the networks possess random structure (as opposed to, e.g. lattice networks on a torus, where all nodes have identical degree but are only connected to their nearest neighbors). The complete graph with N nodes is an $(N-1)$-regular network and thus a special case of k-regular random networks. The cover time of the complete graph is given as $\langle T \rangle = (N-1)\left(\log(N-1) + \gamma + \mathcal{O}(N^{-1})\right)$ [99], a result which is reproduced by Eq. (6.10) since the GMFPT for each node[2] is $\tilde{\tau} = N - 1$ and $\psi(N+1) = \log N + \mathcal{O}(N^{-1})$. For general random k-regular networks, Eq. (6.10) can be used to find an approximate scaling relation for the lower bound

$$\langle T \rangle \gtrsim \frac{k}{k-1} N \log N \tag{6.11}$$

using the GMFPT lower bound Eq. (3.14), the fact that $k_v = \langle k \rangle = k$, and $\psi(N+1) = \log N + \mathcal{O}(N^{-1})$.

[2] Suppose a random walker starts at any node u. The probability to reach any other node of the network in one time step is $p = 1/(N-1)$. Looking at a single target node v the probability that v is first passaged at time t is given as $p_t = (1-p)^{t-1}p$. Hence, the GMFPT for every target node is $\tilde{\tau} = \sum_{t=1}^{\infty} t p_t = -p \frac{\partial}{\partial p} \frac{1}{p} = \frac{1}{p} = N - 1$.

Network	N	$\langle k \rangle$	$\frac{\langle T^{\text{sim}} \rangle}{N \log N}$	$\frac{\langle T^{\text{es}} \rangle}{N \log N}$	rel. err.	$\frac{\langle T^{\text{ex}} \rangle}{N \log N}$	rel. err.
Intra-org. contacts - Cons. (info) [123]	43	15.3	2.40	2.41	0.004	2.38	0.007
Intra-org. contacts - Cons. (value)	44	16.0	2.00	2.02	0.006	2.07	0.031
Intra-org. contacts - Manuf. (awareness)	77	25.5	3.43	3.47	0.012	3.46	0.009
Intra-org. contacts - Manuf. (info)	76	23.3	2.35	2.29	0.028	2.37	0.007
Matches of the NFL 2009 [124]	32	13.2	1.20	1.27	0.053	1.21	0.008
Connections between 500 largest US airports [125]	500	11.9	12.34	10.30	0.198	13.18	0.063
B. F. Maier's Facebook friends network [6]	329	11.9	11.63	8.45	0.377	12.36	0.059
C. Elegans neural network [1]	297	14.6	8.63	9.15	0.057	8.69	0.007
E. Coli protein interaction [126]	329	2.8	5.55	4.27	0.298	7.24	0.233
Social interaction in dolphins [12]	62	5.1	4.72	4.46	0.059	4.86	0.029
American college football [127]	115	10.7	1.38	1.27	0.081	1.40	0.017
Food web of grassland species [128]	75	3.0	4.66	3.97	0.172	5.17	0.099
Zachary's Karate club [85]	34	4.5	3.04	3.29	0.079	3.06	0.006
Interactions in "Les Misérables" [129]	77	6.6	6.81	6.21	0.097	7.21	0.055
Network of associations between terrorists [70]	62	4.9	4.65	4.47	0.040	4.88	0.046

Table 6.1: Mean cover times of simple discrete-time random walks on the largest component of various networks, in units of the cover time on a complete graph with equal node count. Displayed is the number of nodes N, the mean degree $\langle k \rangle$ of the largest component and the measured mean cover time $\langle T^{\text{sim}} \rangle$ extracted from 50 simulations per network with one walker starting on every node. Additionally shown are both theoretical estimations of the cover time using (i) estimated GMFPTs $\langle T^{\text{es}} \rangle$ from the target nodes' degrees, Eq. (6.8) and (ii) exact GMFPTs $\langle T^{\text{ex}} \rangle$ computed from the unnormalized graph Laplacian's spectrum Eq. (6.9). Both estimations are given with their relative error to the simulated mean cover time. Note that all networks have been symmetrized and an unweighted link (u, v) has been created if a weight between two nodes was $w_{uv} > 0$. For the Intra-organizational networks, a link was created if both nodes put anything else than "I do not know this person" in their questionnaire. These values are additionally shown in Fig. 6.5 (top). The table was reproduced from [92], © 2017 American Physical Society.

6.2 *Numerical Results*

In this section, the predictions of Eqs. (6.8) and (6.9) are compared with simulation results for single component ER, BA and real-world networks, as well as Eq. (6.10) for random k-regular networks. For the simulations, a single walker was placed on every node at time $t = 0$. Subsequently, each walker performed a random walk as described in Sec. 3.1.6. Each walker proceeded until it visited each node at least once, completing total coverage and marking cover time T_u. $\langle T \rangle$ was computed as the average of all T_u.

For both ER and BA networks, networks were generated with $N \in \{50, 100, 200, 400\}$ nodes, ER networks with node connection probability $\{k/(N-1) | k \in \mathbb{N},\ 1 \leq k \leq 20\}$, and BA networks with $\{m | m \in \mathbb{N},\ 1 \leq m \leq 20\}$. In order to test Eq. (6.10), random k-regular networks were generated using the algorithm given in [130] with $N \in \{50, 100, 200, 400\}$ nodes and node degree $k_u = k\ \forall\, u \in \mathcal{V}$, scanning integer degrees $\{k | k \in \mathbb{N},\ 3 \leq k \leq 20\}$. Network generation was followed by the extraction of the largest component and a subsequent random walk simulation as described above. The cover time was estimated using Eqs. (6.8), (6.9), and Eq. (6.1), respectively, for 1000 networks each. For Eq. (6.10) and the random k-regular networks, the global mean FPTs $\tilde{\tau}^{\text{es}} = N/(1 - k^{-1})$ and $\tilde{\tau}^{\text{ex}} = N^{-1} \sum_{v=1}^{N} \tau_v^{\text{ex}}$ were used, respectively.

The theoretic results are in agreement with the simulation results, as can be seen in Fig. 6.2. The relative error decreases with increasing number of nodes N as well as increasing mean degree $\langle k \rangle$ and quickly reaches values below 1%. Unsurprisingly, the heuristic method performs better compared to the results of [111, 112] due to the asymptotic nature of the latter.

In addition to simulations on network models, simulations were performed for the largest component of 15 real-world networks, listed in Tab. 6.1. Inititially directed networks were converted to undirected

Network	N	$\langle k \rangle$	$\frac{\langle T^{\text{sim}} \rangle}{N \log N}$	$\frac{\langle T^{\text{es}} \rangle}{N \log N}$	$1 - \frac{\langle T^{\text{sim}} \rangle}{\langle T^{\text{es}} \rangle}$	$\frac{\langle T^{\text{ex}} \rangle}{N \log N}$	$1 - \frac{\langle T^{\text{sim}} \rangle}{\langle T^{\text{ex}} \rangle}$
Barcelona	128	2.2	9.7	2.7	2.6	12.3	0.2
Beijing	104	2.2	10.5	2.7	2.9	16.0	0.3
Berlin	170	2.1	14.7	2.8	4.3	21.2	0.3
Chicago	141	2.1	14.9	2.7	4.4	20.0	0.3
Hong Kong	82	2.1	10.1	3.0	2.4	13.1	0.2
London	266	2.3	14.8	2.8	4.3	19.3	0.2
Madrid	209	2.3	14.4	2.7	4.4	20.4	0.3
Mexico	147	2.2	10.9	2.7	3.0	14.8	0.3
Moscow	134	2.3	12.0	2.9	3.1	14.7	0.2
New York	433	2.2	16.2	2.6	5.1	24.3	0.3
Osaka	108	2.3	8.9	2.9	2.1	11.4	0.2
Paris	299	2.4	11.4	2.8	3.0	14.2	0.2
Seoul	392	2.2	19.8	2.6	6.7	31.2	0.4
Shanghai	148	2.1	14.5	2.7	4.4	20.0	0.3
Tokyo	217	2.4	12.9	2.8	3.7	17.7	0.3

Table 6.2: Same procedure as in Tab. 6.1, but for subway networks of big cities, taken for the year 2009 from [131]. These values are additionally shown in Fig. 6.5 (bottom). The table was reproduced from [92], © 2017 American Physical Society.

networks replacing every directed link with an undirected link. For weighted networks an undirected link (u, v) was assigned if a weight was $w_{uv} > 0$. For the intra-organizational networks [123], employees had to fill out questionnaires regarding their relationships to coworkers. Here, an undirected link (u, v) was assigned if both u and v marked anything else than "I do not know this person". As can be seen in Tab. 6.1, the heuristic method produces results that are satisfyingly close to the simulated values (mostly relative errors of $< 10\%$). Exceptions are the computed cover times for the *E. coli* protein interaction network [126] with a relatively high relative error of $\approx 23\%$ and the grassland food web [128] with a relative error of $\approx 10\%$.

Furthermore, some example estimate cover time distributions as per Eq. (6.6) are compared to measured distributions from the simulation. As can be seen in Fig. 6.4, the cover time distributions using the exact GMFPTs predict the cover time pdf and complementary cdf (ccdf) reasonably well. In some cases, even the estimated GMFPTs show reasonable agreement. The last panel in Fig. 6.4 shows some significant deviation from the simulations, which will be further discussed below.

6.3 Error Analysis

6.3.1 Low-Dimensionality Induced Deviations

Concerning the impact of network structure on the error of the heuristic compared to the true mean cover time, networks with a large mixing time were consistently associated with higher errors. Since the results are derived under the assumption that the relaxation time is $t_{\text{mix}} \ll N$, the relative error was measured against the ratio t_{mix}/N

to find

$$\text{rel. err.} \approx 0.19 \times \left(\frac{t_{\text{mix}}}{N}\right)^{0.94} \tag{6.12}$$

as can be seen in Fig. 6.6. Here, t_{mix} was computed using Eq. (3.7). This result indicates that an increasing relaxation time indeed increases the error of the heuristic. Additionally, simulations were performed on D-dimensional lattices of dimension $D \in \{1, 2, 3\}$ (chains, squares and cubes) using node numbers $N \in \{(2n+1)^2 | n \in \mathbb{N},\ 2 \leq n \leq 12\}$ for $D = 2$ and $N \in \{n^3 | n \in \mathbb{N}, 4 \leq n \leq 8\}$ for $D = 1$ and $D = 3$. For low-dimensional lattice networks with $D \leq 2$, the relaxation time is large compared to a variety of complex networks (see Fig. 1 in [132]). Hence, one might suspect that the method introduced above will not perform well for low-dimensional lattice networks.

Indeed, as can be seen in Fig. 6.3, the relative error between simulation and heuristic results increases with increasing N, up to $\approx$ 110% for chains and $\approx$ 10% for square lattices using exact GMFPTs, whereas smaller relative errors of up to $\approx$ 4% are reached for cube lattices. Similar results are obtained for real-world networks embedded in a two-dimensional space with short-range connection probability such as subway networks [131] (shown in Tab. 6.2 and Fig. 6.5). Here, the estimation from estimated GMFPTs systematically underestimates the cover time while using exact GMFPTs yields an overestimation of the cover time by $\approx 20\% - 40\%$. Generally, the more exact result of GMFPTs calculated via the unnormalized graph Laplacian gives results with lower relative error than using lower bound GMFPTs, as expected.

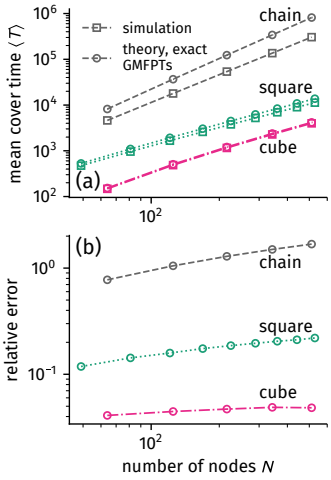


Figure 6.3: Example of the heuristic method yielding results with rather large deviations from simulations. **(a)** Mean cover time for low-dimensional ($D \leq 2$) lattices as well as lattices in dimension $D = 3$ as (□) an average over 1000 simulations for each data point and (◯) theoretical result from exact GMFPTs computed from the unnormalized graph Laplacian's spectrum Eq. (6.9). **(b)** The relative error is increasing with increasing system size but is comparably lowest for $D = 3$ (cubes). The figure was reproduced from [92], © 2017 American Physical Society.

This result can be even more refined by investigating the relative error of the method for networks generated from the small-world models devised in Ch. 4 as those naturally interpolate between low-dimensional lattice-like networks where the heuristic has large deviations and Erdős–Rényi networks where the heuristic yields satisfying results. Such an analysis will be done in Sec. 6.4.

6.3.2 *Systematic Error: Continuous-Time Approximation*

Considering discrete time, the mean cover time is given as the series, respectively partial sum

$$T_u = \sum_{T=1}^{\infty} [1 - P_u(T)] = \sum_{T=1}^{\infty} \bar{P}_u(T) \approx \sum_{T=1}^{T_{\max}} \bar{P}_u(T)$$

where the upper boundary was approximated by numerically finding a $T_{\max}$ with $\bar{P}_u(T_{\max}) \leq 10^{-10}$. This partial sum is equal to the trapezoidal approximation of the integral

$$\int_0^{T_{\max}} \mathrm{d}T\ \bar{P}_u(T) = \frac{1}{2}\bar{P}_u(0) + \sum_{T=1}^{T_{\max}-1} \bar{P}_u(T) + \frac{1}{2}\bar{P}_u(T_{\max}) + \Phi$$

with $\Delta T = 1$. Since the function $\bar{P}_u(T)$ has value $\bar{P}_u(0) = 1$, using the integral instead of the sum introduces a systematic error of

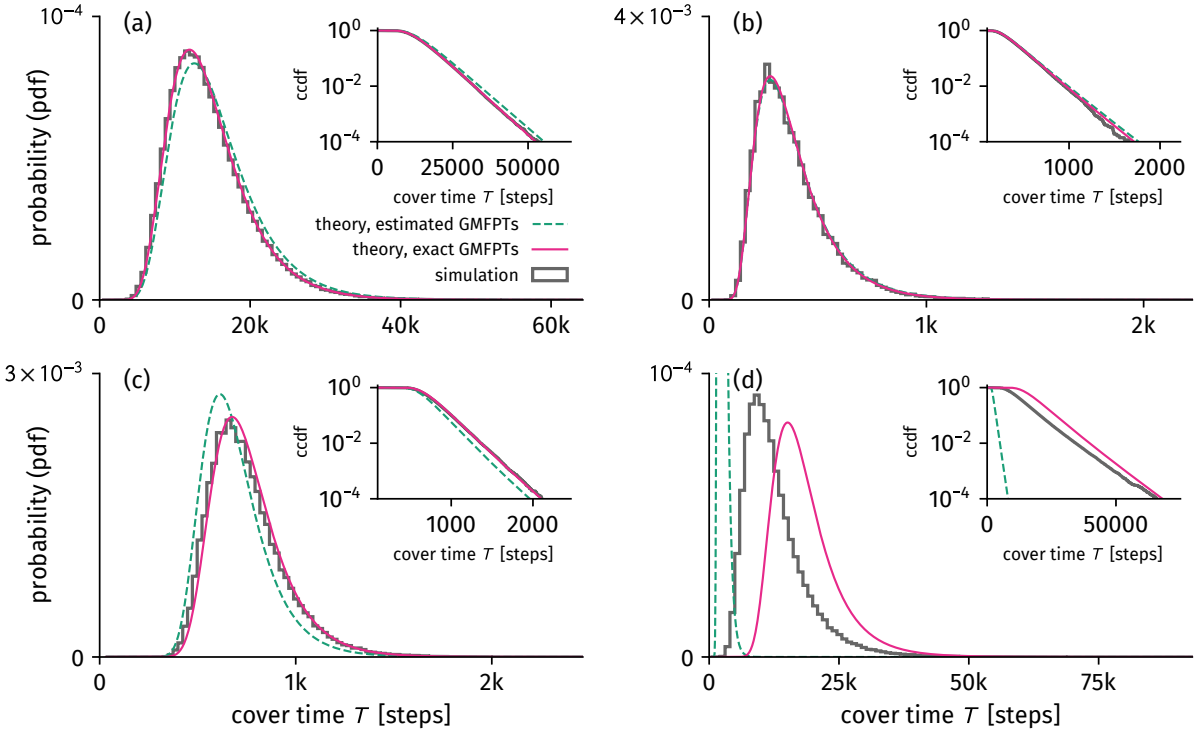


Figure 6.4: Cover time probability distribution functions (pdfs) and complementary cumulative distribution functions (ccdfs, insets) for four example networks **(a)** C. Elegans, **(b)** Intra-org.–Consulting (info), **(c)** College football and **(d)** Berlin's subway system from 2009. Dashed lines show the distributions from estimated arrival rates as per Eq. (6.8) and solid pink lines show the distributions from exact arrival rates (see Eq. (6.9)).

1/2. Using the first derivative $\bar{P}'_u(T)$, the error Φ emerging from the trapezoidal rule can be asymptotically estimated to be

$$|\Phi| = \frac{\Delta T^2}{12} \left| \bar{P}'_u(T_{\max}) - \bar{P}'_u(0) \right|$$

for $T_{\max} \to \infty$ [133]. With

$$\bar{P}'_u(T) = -\sum_{v \neq u} \beta_v \exp(-\beta_v T) \prod_{w \neq v \neq u} [1 - \exp(-\beta_w T)]$$

one has $\bar{P}'_u(0) = 0$. In another way, analogously to Eq. (6.7) one can find

$$\bar{P}_u(T) = \sum_{\mathcal{S} \in \mathcal{P}^*(\mathcal{V} \setminus \{u\})} (-1)^{|\mathcal{S}|+1} \exp\left(-\sum_{v \in \mathcal{S}} \beta_v T \right)$$

$$\left| \bar{P}'_u(T) \right| = \sum_{\mathcal{S} \in \mathcal{P}^*(\mathcal{V} \setminus \{u\})} (-1)^{|\mathcal{S}|+1} \left(\sum_{v \in \mathcal{S}} \beta_v \right) \exp\left(-\sum_{v \in \mathcal{S}} \beta_v T \right).$$

For most nodes the decay rates are $\beta_v \lesssim N^{-1}$ with $\lesssim$ meaning "lower or of similar order". Then $\sum_{v \in \mathcal{S}} \beta_v \lesssim 1$ and hence $|\bar{P}'_u| \lesssim \bar{P}_u$ such that one can safely assume $|\bar{P}'_u(T_{\max})| \lesssim \bar{P}_u(T_{\max}) \leq 10^{-10}$ yielding absolute error

$$|\Phi| \lesssim 10^{-11},$$

which will be sufficiently small for most applications. For applications where this is not small enough, $T_{\max}$ can be further increased until a satisfying error is reached.

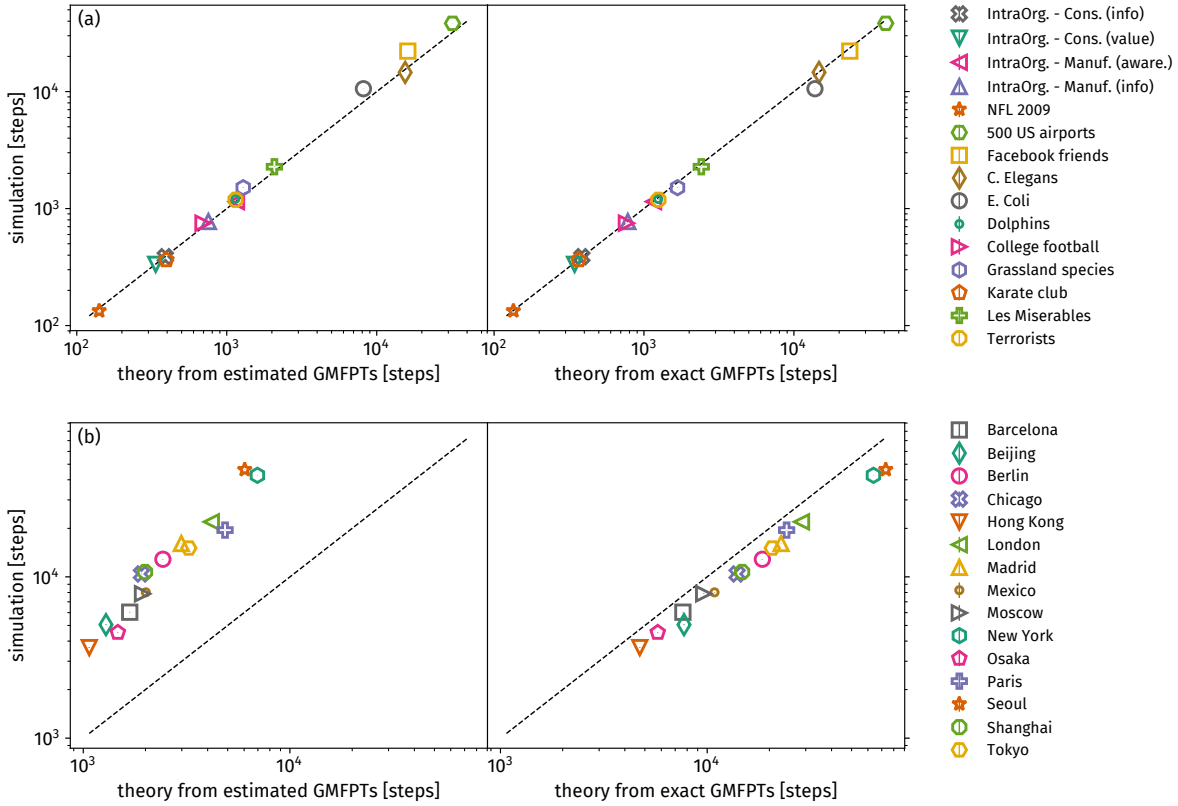


Figure 6.5: Mean cover times of simple discrete time random walks on the largest component of **(a)** various real-world networks (data sources and relative errors given in Tab. 6.1) and **(b)** various subway networks (data source and relative errors given in Tab. 6.2). The estimated cover times are compared to the measured cover times (from 50 simulations for each data point). The dashed lines represent the ideal case $\langle T^{\text{sim}} \rangle = \langle T^{\text{es/ex}} \rangle$. Theoretical results are computed from **(left)** estimated GMFPTs from the target nodes' degrees Eq. (6.8) and **(right)** exact GMFPTs computed from the unnormalized graph Laplacian's spectrum Eq. (6.9). The figure was reproduced from [92], © 2017 American Physical Society.

6.3.3 *Systematic Error: Approximation of Mean Cover Time Integral*

In the following, it is demonstrated that instead of solving integral Eq. (6.5), one can safely use Eq. (6.7). The total difference between both is

$$\begin{aligned}\Theta &= \int_0^\infty \mathrm{d}T\,[1-P(T)] - \\ &\quad - \int_0^\infty \mathrm{d}T\left[1 - P(T)\frac{1}{N}\sum_{v=1}^{N}\frac{1}{1-\exp(-\beta_v T)}\right] \\ &= \int_0^\infty \mathrm{d}T\; P(T)\,[Q(T)-1]\,, \end{aligned} \tag{6.13}$$

defining $Q(T) = \frac{1}{N}\sum_{v=1}^{N}(1-\exp(-\beta_v T))^{-1}$. Note that the cover time cdf $P(T)$ is given by Eq. (6.6), s.t. both

$$\lim_{T\to 0} P(T) = 0$$
$$\lim_{T\to 0} P(T)Q(T) = 0$$

and

$$\lim_{T\to \infty} P(T) = 1$$
$$\lim_{T\to \infty} P(T)Q(T) = 1,$$

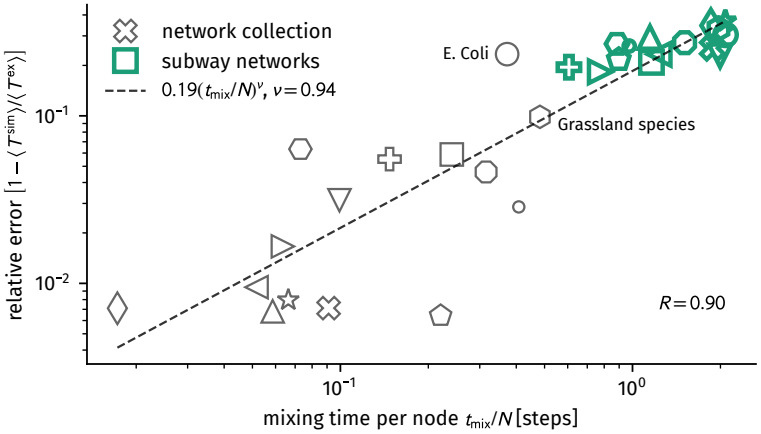


Figure 6.6: The relative error of the mean cover time heuristic is increasing with increasing mixing time per node t_{mix}/N, indicating that the heuristic produces higher deviations for networks with relatively high mixing times as per Eq. (3.7). Networks are marked with the same symbols as in Fig. 6.5. The figure was reproduced from [92], © 2017 American Physical Society.

meaning that for both integration limits, the integrand approaches zero. The following argumentation is based on the assumption that the distribution of decay rates is relatively homogeneous in the region of small rates, implying that there is a low number $1 < n \ll N$ of nodes $i \in \mathcal{V}_{\text{small}}$ with $n = |\mathcal{V}_{\text{small}}|$ that are of the same order as $\beta_{\text{min}} = \min\{\beta_v | v \in \mathcal{V}\}$. This is a relatively safe assumption for most network models and real-world networks as in most cases there are more nodes with small degree (hence small decay rates) than nodes with high degree (hence high decay rates). Now suppose the integration approaches a time where $T \approx \beta_{\text{min}}^{-1}$, implying that, while most terms $1 - \exp(-T\beta_{v\notin\mathcal{V}_{\text{small}}})$ are virtually equal to 1 there are still n terms $1 - \exp(-T\beta_i) < 1$, such that $P(T) \approx \prod_{i\in\mathcal{V}_{\text{small}}}^{n}(1 - \exp(-T\beta_i)) \ll 1$. Furthermore, there will already be a majority of terms $1 - \exp(-T\beta_v) \to 1$ which leads to $Q(T)$ approaching $Q(T) \overset{N\gg1}{\longrightarrow} 1$. Hence, one can safely assume that for a network with a larger number of nodes the integrand approaches zero at all times while the global mean cover time grows quickly and thus the relative error of Eq. (6.7) is approaching

$$\frac{\Theta}{\langle T\rangle} \overset{N\gg1}{\longrightarrow} 0.$$

In particular, one can calculate the error between the integrals for random networks with constant GMFPT $\tilde{\tau}$ for every node, which is given as

$$\begin{aligned}\tilde{\Theta} &= \tilde{\tau}_{N+1}\Big[\gamma + \psi(N+1)\Big] - \tilde{\tau}_N\Big[\gamma + \psi(N)\Big] \\ &\approx \tilde{\tau}_N \log\left(\frac{N}{N-1}\right) \approx \tilde{\tau}_N \frac{1}{N}\end{aligned}$$

where $\psi(N+1) = \log N + \mathcal{O}(N^{-1})$ and $\tilde{\tau}_N \simeq \tilde{\tau}_{N+1}$ was assumed where $A \simeq B$ means $\lim_{N\to\infty} A/B = 1$. Consequently, one finds the relative error to be approximately

$$\frac{\tilde{\Theta}}{\langle T\rangle} \approx \frac{1}{N\gamma + N\log(N-1)} \simeq \frac{1}{N\log N}. \tag{6.14}$$

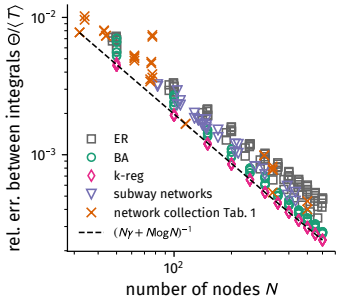


Figure 6.7: Relative error Eq. (6.13) between integrals Eq. (6.5) and Eq. (6.7) for all networks investigated in this chapter besides lattices and the modular hierarchical network models. For each network the relative error for both sets of rates, τ_v^{es} and τ_v^{ex}, is shown. For ER, BA and random k-regular networks, means over the largest components of 100 network realizations for $\langle k\rangle \in \{3,5,7,9\}$ and $m \in \{3,5,7,9\}$ were built, respectively. The measured relative errors are roughly following the scaling relation Eq. (6.14). The figure was reproduced from [92], © 2017 American Physical Society.

Even though this relation is derived for the special case of networks where every node has the same GMFPT, a numerical analysis of Eq. (6.13) reveals that this scaling relation holds approximately for all networks investigated in this chapter, as can be seen in Fig. (6.7).

6.4 Cover Time on Modular Hierarchical Small-World Models

The purpose of studying the cover time on the small-world models introduced in Ch. 4 in the following is two-fold.

First, it raises the opportunity to test the hypothesis that besides the pair-averaged first passage time, which is an upper bound for search times, the mean cover time is minimal between the highly clustered and the random regimes of the models, as well—an upper bound for *exhaustive* search processes. After the findings of Ch. 5, this would render some intermediate modular hierarchical structure optimal for an additional process.

Second, it allows for a more detailed error investigation on when the heuristic begins to fail to predict the cover time accurately as those structures represent a continuous interpolation between the network structures investigated in Secs. 6.2 and 6.3.1.

6.4.1 Average Medium Approximation

As has been shown in Sec. 5.1.1, an average medium approximation (AMA), where a possible edge is replaced with the probability of this edge existing, may provide useful insights into the general mechanisms of random walks on modular hierarchical small-world networks as it allows for thorough analytic considerations. Indeed, one consequence of the AMA is that every node has the exact same GMFPT Eq. (5.5) which allows using Eq. (6.10) to estimate the mean cover time.

In order to test the validity of Eq. (6.10) for average medium self-similar small-world (SSMH) networks, random walk simulations have been performed on averaged SSMH networks and averaged 1D lattice power-law small-world (PLSW) networks; the same simulations as in Sec. 5.1.1. For $B = 8$, $L = 3$ (number of nodes $N = 8^3$) and $\langle k \rangle = 7$, a single walker was placed on each node of each average medium (SSMH and PLSW). Random walk simulations were performed as described in Sec. 3.1.1, where for each walker the simulation was stopped as soon as it visited each node at least once, marking the cover time for this walker, subsequently computing the mean to obtain the mean cover time. The results can be seen in Fig. 6.8. The heuristic Eq. (6.10) approximates the simulated cover time reasonably well for moderate degrees of hierarchical clustering and starts to deviate more strongly from the simulation results for structural control parameters $\xi \lesssim 0.25$. This is not surprising because the heuristic evaluation of the mean cover time is based on the assumption that the mixing time is small which is violated for

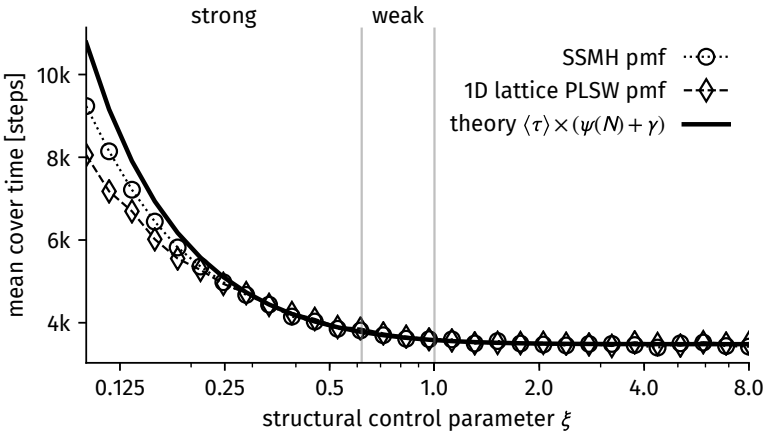


Figure 6.8: The mean cover time as estimated with an average medium approximation of self-simular modular hierarchical as well as 1D lattice PLSW networks as described in Sec. 5.1.1. Shown are simulation results on the averaged media as well as Eq. (6.10) with $\tilde{\tau}$ as given by Eq. (5.6). The figure was reproduced from [93].

strongly modular hierarchically clustered networks (see Sec. 5.1.1).

The result of the AMA is hence that no optimal mean cover time exists between the highly regular structures and the random case.

6.4.2 *Actual Network Realizations*

Here, the mean cover time is analyzed for actual network samples of the self-similar modular hierarchical (SSMH), the modified Watts–Strogatz (mod. WS), the one-dimensional lattice power-law small world (1D latt. PLSW) and the 2-torus PLSW models. To this end, networks with parameters $B = 8$, $L = 3$ (corresponding to $N = 8^3$) and $\langle k \rangle = 8$ where sampled for varying values of their respective control parameters. The control parameters of the SSMH, the mod WS and the 1D latt. PLSW model were mapped using Eq. (4.29) and Eq. (4.27). For each parameter combination, 740 independent network samples were drawn, subsequently their largest component was extracted to be used for analysis. For each largest component, 100 random walkers were placed on 100 random nodes. Random walk simulations as described in Sec. 5.2.1 were started and run until each walker visited each node at least once. When a walker reached its last remaining non-visited node, this time step was saved as its cover time. The mean cover time of that network sample was then computed as the mean of those saved times. The theoretical mean cover time of that network sample was computed using Eq. (6.9). Subsequently, the respective means were built over the 740 obtained single-network means.

The results are shown in Fig. 6.9. Since Eq. (6.9) relies on the global mean FPT which is responsible for a minimum in the pair-averaged FPT, it is not surprising that the mean cover time reveals a minimum, as well (for the SSMH, mod. WS and 1D latt. PLSW models). The position of this optimum in exhaustive search time lies, however, in the strong hierarchically clustered regime where average nodes are connected to more nodes of short-range distance than to any nodes positioned at long-range distance. This emergence of the

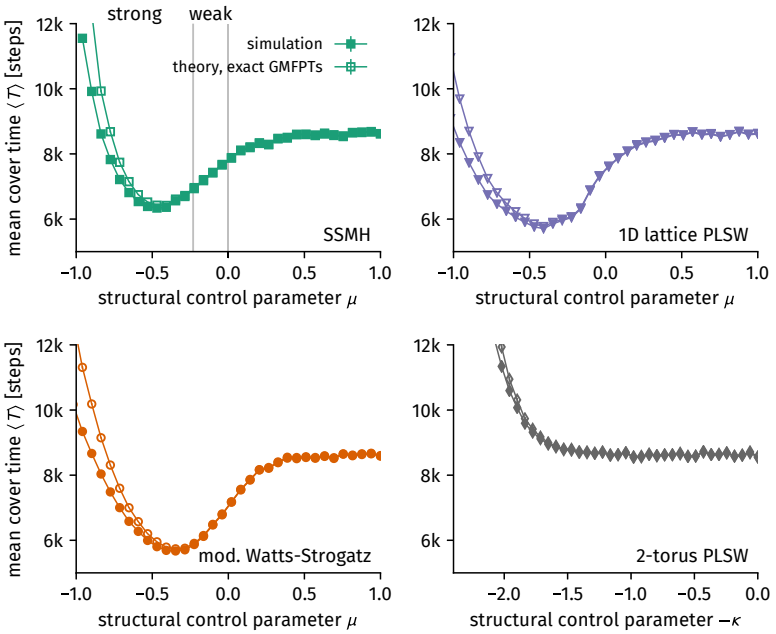


Figure 6.9: The mean cover time on actual network realizations of the (modular hierarchical) small-world models introduced in Ch. 4. Parameters $B = 8$, $L = 3$, and $\langle k \rangle = 8$ were used, markers denote averages over 740 independent network realizations. For each network realization, the cover time was measured from simulations as well as computed from Eq. (6.9). All network models based on discrete node positions display minimal mean cover time in the strong modular hierarchically clustered regime since those display increasing node degree variance with increasing long-range connection probability. The network model based on random node positions (2-torus power-law small-world model) does not as it has constant node degree variance over the full range of its structural control parameter.

minimum can be explained analogously to the argumentation given in Sec. 5.2.2. For high clustering (and low dimensionality), the cover time becomes large but decreases with increasing long-range connection probability and concurrent decreasing clustering. At the same time the node degree variance increases such that more and more nodes of small degree appear which take longer to visit and hence increase the cover time. This second effect does not contribute to the cover time if the degree distribution is constant when increasing long-range connection probability as is the case for the 2-torus PLSW model. Consequently, a minimum in cover time is not observed for network samples of this model.

As implied, the analyses of this section can also be used to obtain an improved intuition about the error produced by the heuristic evaluation of the cover time. The relative error between simulation results and Eq. (6.9) is shown in Fig. 6.10. As expected, the relative error systematically increases with decreasing structural control parameters as soon as values are reached which produce structures being associated with modular hierarchical clustering. For real-world networks there is no such thing as a measurable structural control parameter, and hence, the mixing time per node t_{mix}/N will be used as a proxy for the amount of 'locality' of the networks. For all network samples used in the procedure described above, the mixing time was found using Eq. (3.7) and subsequently averaged. As can be seen in the top panel of Fig. 6.10, the functional dependence of the relative error does not correspond ideally to a power-law (with decreasing structural control parameter (i.e. increasing t_{mix}/N), the relative error flattens off). Nevertheless, the approximate relation

$$\text{rel. err.} \approx 0.81 \times \left(\frac{t_{\text{mix}}}{N}\right)^{1.17} \tag{6.15}$$

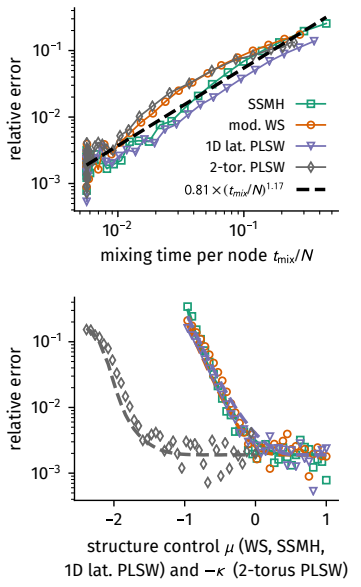


Figure 6.10: Relative errors between simulation results and Eq. (6.9) for the analysis displayed in Fig. 6.9. **(Top)** The relative error as a function of the mixing time per node as computed using Eq. (3.7) for each network sample and subsequent averaging. The dashed line represents the fit Eq. (6.15). **(Bottom)** The relative error as a function of the structural control parameters. The dashed line represents the fit Eq. (6.15) using the respective values of t_{mix}/N used in the top panel.

satisfyingly estimates the relative error of the heuristic method (see bottom panel of Fig. 6.10). As this estimation is greater than Eq. (6.12) for larger values of t_{mix}/N it might be more reliable to estimate the relative error using Eq. (6.15).

6.5 *Summary and Discussion*

A new heuristic method to find the mean cover time of random walks on arbitrary complex networks has been introduced and discussed. It is based on the assumption that arrival times follow an exponential distribution whose mean is given by the respective target node global mean first passage time. The heuristic was applied to several synthetic networks as well as real-world networks. It was found that the heuristic results become increasingly unreliable when their mixing time becomes increasingly large. For these cases the approximation of exponentially distributed arrival times is only valid asymptotically, the estimated arrival times are consequently wrongly estimated which in turn influences the mean cover time. Ways to estimate the relative error of the method were given, however.

The heuristic was subsequently applied to the modular hierarchical network models defined in Ch. 4. As for the pair-averaged first passage time of Ch. 5 minima in the mean cover time were found for all network models where node degree variance increases with increasing long-range connection probability. The remaining modular hierarchical network model where node degree variance is constant does not display a minimum in mean cover time.

Therefore one cannot state that in general, modular hierarchical networks provide an optimal structure for fast uninformed exhaustive searches.

Part III

Epidemics on Temporal Face-to-Face Contact Networks

7
Flockworks: Analyzing Temporal Face-to-Face Contact Networks With a Simple Model

The third part of this thesis investigates how the temporally resolved description of physical contacts between humans influences the spread of diseases in the population. To this end, a reference model which replicates certain network properties is a useful tool to find out which features are responsible for facilitating or mitigating an infection's spread. As discussed in Sec. 2.3, empirical face-to-face contact networks display the following basic properties:

1. At any time, the network consists of disconnected components, or groups, to which nodes connect or from which they disconnect.
2. Inter-contact durations, contact durations, and group life-times[1] follow heavy-tailed distributions. The shape of the group life-time distribution does not change with group size but the maximal group-life time increases with increasing group size.
3. Network properties follow a circadian activity rhythm.

[1] For definitions, see Sec. 2.3.2.

Temporal networks generated by a model introduced in Sec. 2.4.1 are associated with properties 1-2. However, nodes group structure and heterogeneous interaction times were postulated and do not emerge from underlying principles in these models. They further implement heterogeneous interaction times by introducing memory effects while ignoring the circadian nature of real-world networks. This raises the question whether simpler models exist for which the network properties 1-3 emerge from first principles. In particular, heterogeneously distributed inter-event times may be explained by a Poissonian model of varying rates, as was shown for interaction times of e-mail communication [47, 48].

To this end, a temporal network model called *Flockwork*[2] will be devised in this chapter. The emergence of the real-world network properties described above will be discussed while the model's accuracy is increased step by step. The model is based on two node behaviors, each associated with a rate. Following the first behavior, a node actively establishes new connections with another node and its neighborhood with *active reconnection rate* α. Second, nodes

[2] Note that the model was developed in cooperation with D. Brockmann who also gave it its name.

may actively decide to stay alone, with rate β, called the *active disconnection rate*. Following these two events, it will be shown that the system always reaches a state in which the network consists of disconnected components which are completely connected within. The size distribution of this state will be stable. However, due to the Poissonian nature of the model, none of the temporal distributions will be heavy-tailed.

In an attempt to increase the realism of the model, temporally varying event rates will be introduced which will lead to a more accurate description of the structural distributions but still yields distributions without heavy tails for the temporal observables. Concerning the circadian nature of the real-world network, it will further be shown that human temporal contact networks might be interpreted to follow a trajectory of equilibrium states in a two-dimensional state space with one structural and one temporal dimension. An inference method for these trajectories is proposed.

7.1 The Poissonian Flockwork Model

Here, the basic version of the Flockwork model will be defined. Subsequently, it will be shown that this model always produces an equilibrium state where the network consists of several disconnected components, or groups, which are all complete networks. The size distribution of these groups is computed, along with asymptotic formulas for its first and second moment, the mean number of components, the degree distribution, the asymptotic mean degree and the group life-times.

7.1.1 Model Definition

Given any network at time t_0, each node can undergo two events.

1. With active reconnection rate α, the node cuts all its links, connects to another node v and connects to all of v's neighbors $v' \in \mathrm{Nei}(v)$.

2. With active disconnection rate β, the node cuts all its links and stays alone.

These rates are meant to reflect basic human behavior in the sense that people may decide not to have any contacts, or decide to establish a new face-to-face contact, which makes it likely to be in contact with all the other people who are already in close proximity to the new contact.

Since in both events all links are cut first, these rates can be mapped to a total event rate $\gamma = \alpha + \beta$. Then the model is described as follows. Each node cuts all its links with rate γ. With probability P it reconnects to any node v and all of v's neighbors. Both descriptions are equivalent. In this section, the γ-P-model will be used because the time scale can be fixed using γ and subsequently the *reconnection probability* P is used as a single parameter to control the structure.

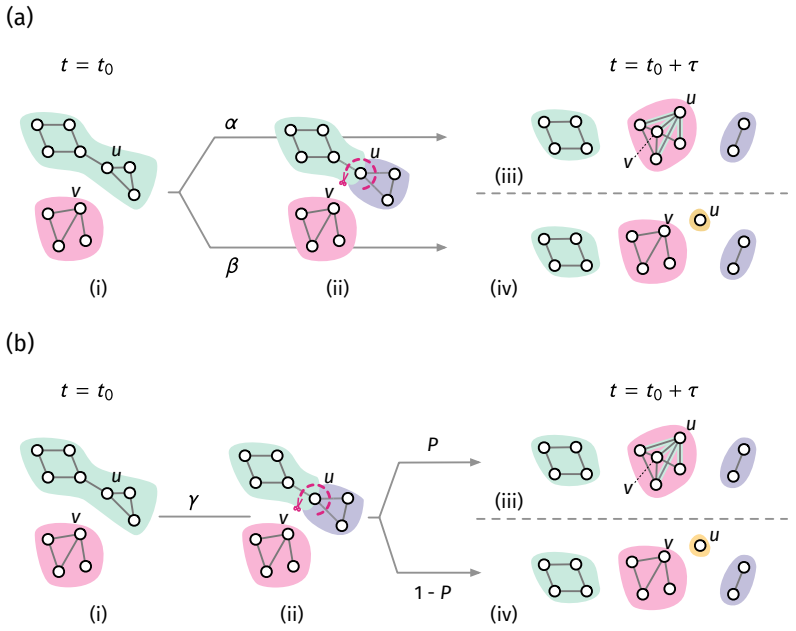


Figure 7.1: The possibile events in a Flockwork model for any network state at $t = t_0$. **(a)** In the α-β-description of the model, each node undergoes a reconnection event with rate α and a disconnection event with rate β. (a.ii) In both cases, the link first cuts all its connections. (a.iii) For an α event it subsequently connects to a random node v and to all of v's neighbors. (a.iv) For a β-event it stays alone. The updated time is then $t = t_0 + \tau$ where τ follows an exponential distribution $\tau \sim \mathcal{E}(\alpha + \beta)$. **(b)** in the γ-P-description, each node undergoes a (b.ii) link cutting event with rate γ. Afterwards, it (b.iii) reconnects with probability P to a random node v and all of v's neighbors. (b.iv) With probability $1 - P$ it stays alone. The updated time is then $t = t_0 + \tau$ where τ follows an exponential distribution $\tau \sim \mathcal{E}(\gamma)$. Both (a) and (b) are equivalent. Both variants are introduced here because both will be used throughout this chapter.

For later descriptions, it is useful to provide mapping equations of the control parameters, which are given as

$$\begin{aligned}\gamma &= \alpha + \beta\,,\\ P &= \alpha/(\alpha+\beta)\,,\\ \alpha &= \gamma P\,,\\ \beta &= \gamma(1-P)\,.\end{aligned}$$

In principle, the number of nodes N in combination with any two of the parameters above define exactly one Flockwork configuration. An illustration of the two descriptions is given in Fig. 7.1.

A simulation method is based on Gillespie's stochastic simulation algorithm as described in App. D. Consider any network at time $t = t_0$. An inter-event time τ is sampled from the exponential distribution $\tau \sim \mathcal{E}(\gamma)$. Afterwards, a node u is picked uniform at random and all edges connecting to node u are deleted from the network. Subsequently, the following happens with probability P: A second node v is chosen uniform at random from the remaining $N - 1$ nodes. Then, edges between u and v as well as u and all $v' \in \mathrm{Nei}(v)$ are added to the network. The time is then updated as $t = t_0 + \tau$.

7.1.2 *Group Sizes and Equilibrium State*

In this section it will first be shown that given a collection of fully connected groups as an initial state, the Flockwork process will approach an equilibrium configuration where the network consists of disconnected components, each of which is fully connected. The group-size distribution is shown to be constant for constant reconnection probability. The discussion is subsequently extended to an

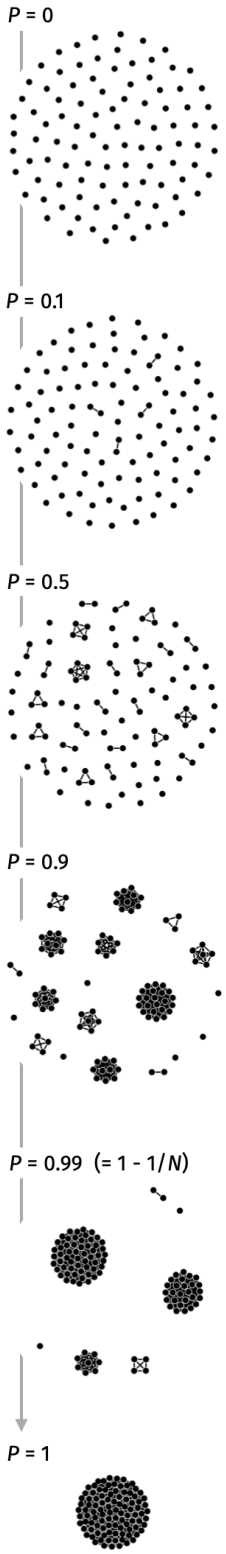


Figure 7.2: Flockwork equilibrium configurations for increasing reconnection probability P (for the networks shown here, $N = 100$ was used). The structural control parameter P directly controls the heterogeneity of the group-size distribution (cf. Fig 7.5).

arbitrary initial condition (initial network) which is shown to approach the same equilibrium configuration. The group-size distribution, mean group size, and mean number of components will be computed analytically. An algorithm is given to construct equilibrium configurations without the need of simulating until an equilibrium state is reached.

Temporal Evolution

Consider an empty network of N nodes as an initial condition. Any event a node undergoes will lead to the network consisting of fully connected groups: Isolated nodes will connect to build pairs, then triangles, then fully connected groups of four and so forth. Whenever a node disconnects, it will stay alone, representing a complete group of $N = 1$ nodes, whereas its former group will remain a completely connected component, albeit consisting of one node less than before. Therefore, it suffices to study the group-size distribution of the Flockwork model to obtain a complete picture of the networks the Flockwork model produces over time.

Denote by n_g the number of fully connected groups of size $1 \leq g \leq N$ such that the N-sized column vector $\boldsymbol{n}$ contains the complete information of the network state. Note that at any time one has $n_0 = n_{N+1} = 0$, which will be used as boundary conditions below—therefore, these are not part of the state vector n_g. Given that the network is initially in a state which is fully defined by the vector $\boldsymbol{n}$, changes in group sizes by the events of the Flockwork dynamic process are governed by the following ordinary differential equations (ODEs):

$$\begin{aligned}
\gamma^{-1}\partial_t n_1 &= -\frac{2P}{N}n_1 + (1-P)\frac{N-n_1}{N} + \\
&\quad + 2\frac{1-P}{N}n_2 + 2\frac{P}{N}\frac{N-2}{N-1}n_2\,, && \text{(7.1a)}\\
\gamma^{-1}\partial_t n_{g>1} &= \frac{P}{N(N-1)}\Big[-2gn_g(N-g) + \\
&\quad + n_{g-1}\left(N(g-1)-(g-1)^2\right) + \\
&\quad + n_{g+1}\left(N(g+1)-(g+1)^2\right)\Big] - \\
&\quad - (1-P)\frac{g}{N}n_g + (1-P)\frac{g+1}{N}n_{g+1}, && \text{(7.1b)}
\end{aligned}$$

as derived in App. C.1. Here, the time scale is fixed by the event rate per node γ and therefore the equations describe a single event. Since disconnections affect each group size equally, the change equations are (i) linear and (ii) imply that groups of size g can either be reduced or increased by one node. Hence, the ODE for the number of groups of size g only depends on its own group count (groups of size g), the number of groups of size $g-1$, and the number of groups of size $g+1$. The count of groups of size $g = 1$ can increase by considering pure disconnection events where one of the $(N - n_1)$ nodes in groups of

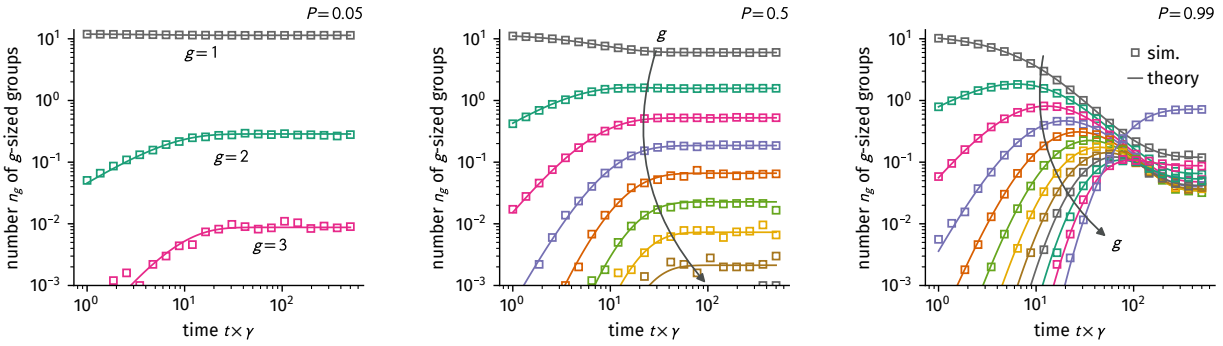


Figure 7.3: The temporal evolution of the expected number n_g of g-sized groups in the Flockwork model for $N = 12$ nodes and reconnection probability **(left)** $P = 0.05$, **(center)** $P = 0.5$, and **(right)** $P = 0.99$. The initial state was an empty network (i.e. $\boldsymbol{n}(0) = (N, 0, \ldots, 0)^T$) and the event rate was chosen as $\gamma = 1$. For each value of the reconnection probability, 10^4 simulations have been performed using the events described in Sec. 7.1.1. Between each event, an inter-event time τ was sampled from the exponential distribution $\tau \sim \mathcal{E}(\gamma)$. Subsequently the group-size distribution n_g was measured which stayed constant until the next event. Markers represent means over the respective number n_g at definite time points in between $t = \gamma^{-1}$ and $t = 500\gamma^{-1}$. Furthermore, Eqs. (7.1) were integrated using a Runge–Kutta solver of order 4 with adaptive stepsize control (Dormand–Prince method). The results, represented by solid lines, are in agreement with the simulations.

size $g > 1$ has a disconnection event and stays alone. The solution of Eqs. (7.1) is illustrated and compared to simulation results in Fig. 7.3 for different values of P.

Since there is no dependence on states other than the current configuration, the Flockwork process is a Markov process. Due to the equations being linear one may build the expectation value of these equations of change and so n_g will be considered to denote the expected number of groups $\langle n_g \rangle$ in the following.

Equilibrium State

For every initial condition where the structure of a network is fully captured by the group count vector $\boldsymbol{n}(0)$, this linear process has a single stationary state $\boldsymbol{n}^\star$ which only depends on the number of nodes N and the reconnection probability P. It is derived in App. C as

$$n_1^\star = N(1-P) \tag{7.2a}$$

$$n_{1<g<N}^\star = \frac{1}{g}\frac{N!}{(N-g)!}\frac{1-P}{\prod_{j=1}^{g-1}[1-Pj/(N-1)]}\left(\frac{P}{N-1}\right)^{g-1} \tag{7.2b}$$

$$n_N^\star = (N-2)!\frac{P}{\prod_{j=1}^{N-2}[1-Pj/(N-1)]}\left(\frac{P}{N-1}\right)^{N-1}. \tag{7.2c}$$

In App. C this result is further extended to be valid for *any* initial network. Any non-fully connected group of size g can transition to produce both fully connected groups *and* non-fully connected groups through the Flockwork dis- and reconnection events (this includes any connected network of size N). However, no fully connected group of size g can produce any non-fully connected group. Hence, after a sufficient amount of time, any initial condition will transition to a network composed of only completely connected groups. Then, Eqs. (7.1) are applicable and hence the equilibrium state Eq. (7.2) is approached. An illustration of this effect is shown for example simulations of $N = 100$ and $P = 0.5$ with two different initial conditions in Fig. 7.4. Therefore, the network fragmentation process to nodes only being in groups or isolated is an emergent property, a consequence of the introduced node behaviors.

Some example equilibrium configurations are shown in Fig. 7.2

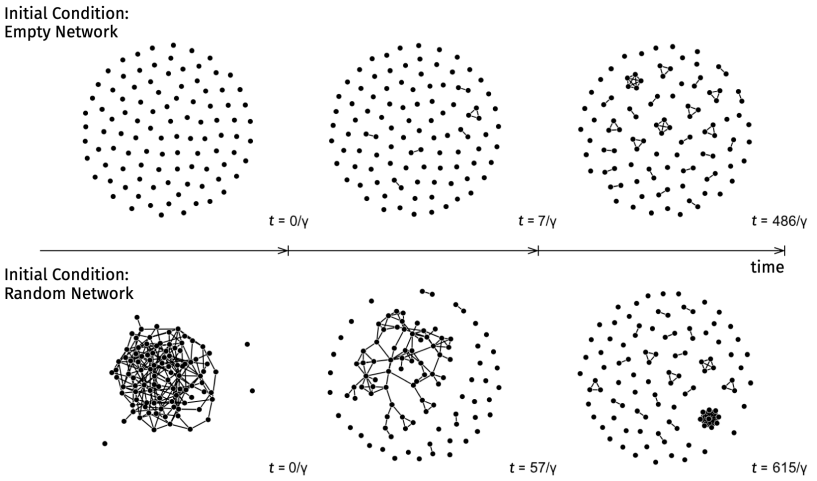


Figure 7.4: Configuration examples for the temporal evolution of a Flockwork model of $N = 100$ nodes and reconnection probability $P = 0.5$ with different initial conditions. **(Top)** the initial state is the empty network. Small groups start to build until a configuration from the equilibrium is reached. **(Bottom)** the initial state is an Erdős–Rényi random network with $\langle k \rangle = 4$. Following the Flockwork process, the network fragments into smaller groups until a configuration from the equilibrium is reached.

for increasing P. Various example group-size distributions are displayed in Fig. 7.5. Note that depending on the reconnection probability P, the number of edges present in the network changes. With $P = 0$, the equilibrium state always corresponds to the empty network, whereas for $P = 1$, the equilibrium state will always be a completely connected network (which is both confirmed by considering these values of the reconnection probability in Eqs. (7.2) which yields $n_1^\star(P = 0) = N$, $n_{g>1}^\star(P = 0) = 0$, $n_N^\star(P = 1) = 1$, and $n_{g<N}^\star(P = 1) = 0$). Hence, the Flockwork model interpolates between those two well-known structures. Also, both of those limit states are static. Therefore even though the model is dynamic, its limiting cases are static networks.

For increasing values of the reconnection probability $0 < P < 1$, the equilibrium group-size distribution changes. It corresponds to a peaked distribution at $g = 1$ for $P = 0$, becoming broader but decaying quickly with increasing group size for smaller values of P until it reaches the distribution

$$n_g^\star(P = 1 - 1/N) = \frac{1}{g}$$

at $P = 1 - 1/N$, as is confirmed by using this value of P in Eqs. (7.2). This will be of further interest below, as empirical networks display heavy-tailed group-size distributions. Further increasing P, the expected peak at $g = N$ develops.

When simulating dynamic systems, it is important that the substrate is in a dynamic equilibrium. Depending on the initial conditions and the parameter choices, it might take a considerable amount of time to reach such an equilibrium configuration when simulating a Flockwork model. However, based on the analytically evaluated equilibrium group-size distribution, equilibrium configurations can be sampled directly. To this end, an efficient algorithm is presented in App. C.2. This algorithm will be used several times in the following sections and the following chapter.

In order to quantitatively describe the group-size distribution with a single observable, the first moment of the group-size distribution

is computed below. The mean group size, denoted as $\langle g \rangle \equiv \langle |\kappa(u)| \rangle$ where $\kappa(u)$ is the set of nodes which are in the same group as u, can be evaluated after the group-size distribution has been properly normalized as

$$\langle g \rangle = \frac{\sum_{g=1}^{N} g n_g^\star}{\sum_{g=1}^{N} n_g^\star} = \frac{\sum_{g=1}^{N} g n_g^\star}{\langle c \rangle} = \frac{N}{\langle c \rangle},$$

where the mean number of disconnected components is given as

$$\langle c \rangle = \sum_{g=1}^{N} n_g^\star.$$

As shown in App. C.1, an exact formula for this quantity can be derived. Asymptotically, one finds

$$\lim_{N\to\infty} \frac{\langle c \rangle}{N} = -\frac{1-P}{P} \log(1-P), \tag{7.3}$$

and

$$\lim_{N\to\infty} \langle g \rangle = -\frac{P}{(1-P)\log(1-P)}. \tag{7.4}$$

Furthermore, the second moment of the group-size distribution is given by

$$\lim_{N\to\infty} \left\langle g^2 \right\rangle = \frac{\langle g \rangle}{1-P}, \tag{7.5}$$

which will be motivated when evaluating the degree distribution in Sec. 7.1.3. These expressions and their corresponding exact values are displayed in Fig. 7.6, showing that the asymptotics are valid for $P < 1 - 1/N$ which is the value at which the group-size distribution approaches a power-law. The asymptotic results can be of importance when a parameter P has to be chosen according to given first and/or second moment of the group-size distribution.

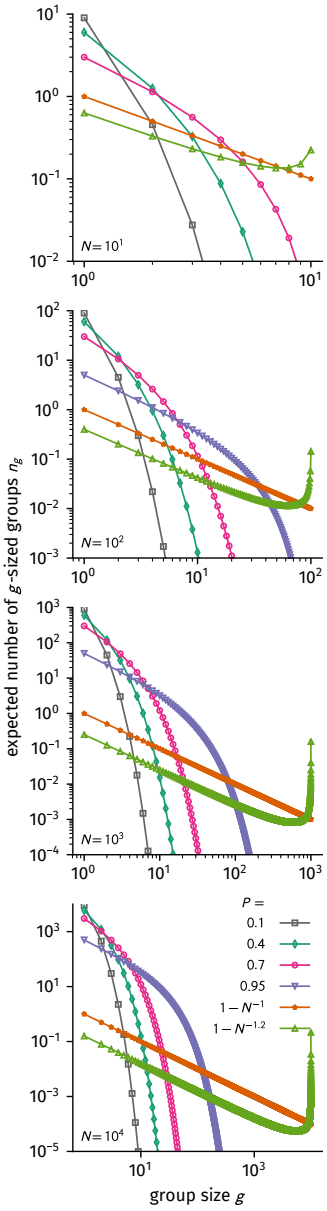
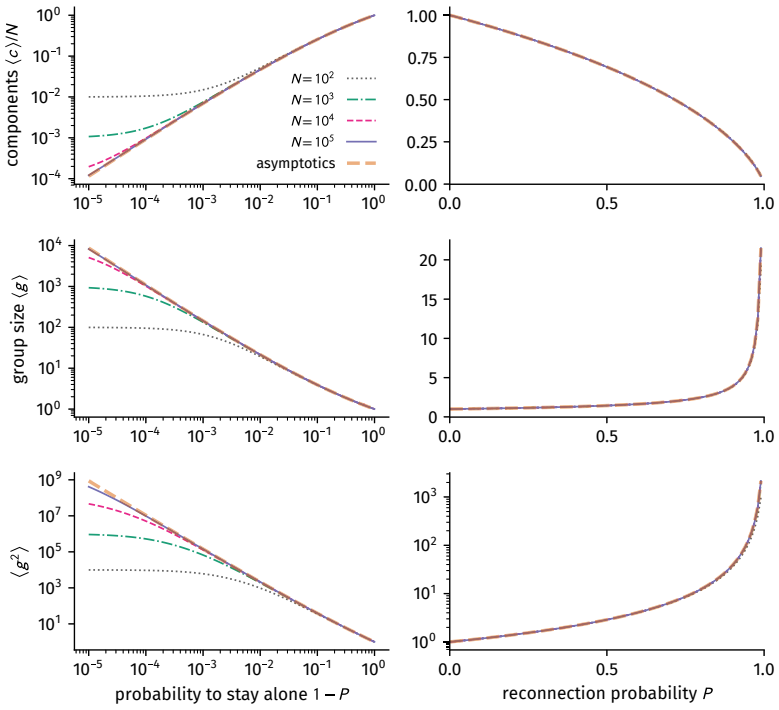


Figure 7.5: Flockwork equilibrium group-size distributions as computed from Eqs. (7.2). Note that the group-size distribution becomes a power-law $n_g^\star = g^{-1}$ for $P = 1 - 1/N$.

7.1.3 *Node Degrees*

As discussed in Sec. 2.1.2, the node degree is an important simple quantity describing how connected a single node is in the network. The mean node degree $\langle k \rangle$ can influence both random walks and epidemic spreading; a heterogeneous degree distribution can cause the epidemic threshold to be shifted towards lower infection rates and hence be responsible for a more susceptible system. Hence, the degree distribution and mean degree of the Flockwork model will be derived in the following, which will be of use later on.

Figure 7.6: Mean group size, mean number of components, and second moment of the equilibrium group-size distribution, all computed from Eqs. (7.2) and compared to the asymptotic Eqs. (7.3-7.5).



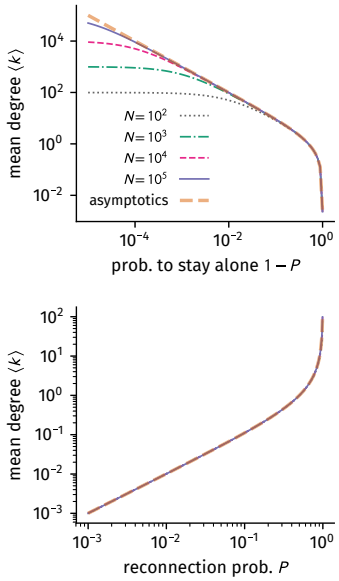


Figure 7.7: Flockwork equilibrium mean degree as a function of the reconnection probability P and number of nodes N, compared to the asymptotics $\langle k \rangle = P/(1-P)$ of Eq. (7.8).

Since the Flockwork model always approaches a configuration where it consists of completely connected groups, each node in a group of size g will have degree $k = g - 1$. The probability of a randomly chosen node having degree k is hence equal to the probability of finding a randomly chosen node in a group of size $g = k + 1$ which is

$$P_k = P[|\kappa(u)| = k + 1] = \frac{k+1}{N} n_{k+1}. \tag{7.6}$$

Hence, the mean degree is given as

$$\langle k \rangle = \frac{1}{N} \sum_{k=0}^{N-1} k(k+1) n_{k+1} = \frac{\langle g^2 \rangle}{\langle g \rangle} - 1, \tag{7.7}$$

see Eq. (C.7) for a detailed derivation.

Furthermore, the mean degree can be computed considering how the expected number $\langle e(t) \rangle$ of edges present in the network changes per event. On average, and considering $N \to \infty$, a γ-event will delete $\langle k(t) \rangle$ edges at time t as an average node cuts all its contacts. Furthermore, with probability P, new edges will be built by connecting to a node of mean degree $\langle k(t) \rangle$ and all its neighbors, thus building $\langle k(t) \rangle + 1$ new edges. Therefore, the temporal evolution of the present number of edges at time t can be approximated as

$$\langle e(t+1) \rangle - \langle e(t) \rangle = - \langle k(t) \rangle + P + P \langle k(t) \rangle ,$$

which asymptotically (with $\langle e(t+1) \rangle = \langle e(t) \rangle$) yields the dynamic equilibrium mean degree

$$\langle k \rangle = \frac{P}{1-P}. \tag{7.8}$$

Combining Eqs. (7.7) and (7.8) gives the asymptotic formula of the second moment of the group-size distribution Eq. (7.5). Example degree distributions are displayed in Fig. 7.8 and the validity of the asymptotic mean degree for $P \lesssim 1 - 1/N$ is confirmed in Fig. 7.7.

7.1.4 *Distributions of (Inter-)Contact Durations and Group Life-Times*

As in this part of the dissertation *temporal* networks are the focus of attention, this section will discuss the basic model's implications for various time-scales, i.e. the inter-contact time, the contact life-time, and the group life-times for a group of size g.

The inter-contact time is the time a node will stay alone before establishing a new contact. The rate with which an isolated node will establish a new connection in the Flockwork model is $2\gamma P = 2\alpha$ (accounting for both an event where the node acts actively as well as an event where it gets passively connected to). Hence, the inter-contact time is exponentially distributed as

$$p(\tau_{IC}) = 2\alpha \exp(-2\alpha\tau_{IC}). \tag{7.9}$$

The group life-times can be evaluated in a similar manner. Suppose a group is of size $g > 1$. Then, three events can lead to a change in group size.

1. One of the g nodes disconnects and stays alone, which happens with rate $g\gamma(1-P)$.
2. One of the g nodes disconnects and reconnects to a node which is not part of the rest of this group, which happens with rate $g\gamma P(1-(g-1)/(N-1))$. Note that it might be more intuitive to express the fraction of the connecting nodes as $(1-(g-1)/(N-1)) = (N-g)/(N-1)$.
3. One of the $N-g$ nodes reconnects to one of the g nodes, which happens with rate $(N-g)\gamma P g/(N-1)$.

The total rate of change for a g-sized group is thus given as

$$\lambda_g = g\gamma(1-P) + 2\gamma P g \frac{N-g}{N-1}$$

and hence the life-time distribution of groups of size g follows as

$$p(\tau_g) = \lambda_g \exp(-\lambda_g \tau_g). \tag{7.10}$$

The contact decay rate can be estimated as follows. Suppose an edge between two nodes has been established. The nodes are in a group of size $g = k+1$ and so they have degree k. Now, two events can influence the life-time of the contact:

1. One of both nodes disconnects. This happens with rate $2\gamma(1-P)$.
2. One of both nodes reconnects to a node which did not belong to its group, which happens with rate $2\gamma P[1-k/(N-1)]$.

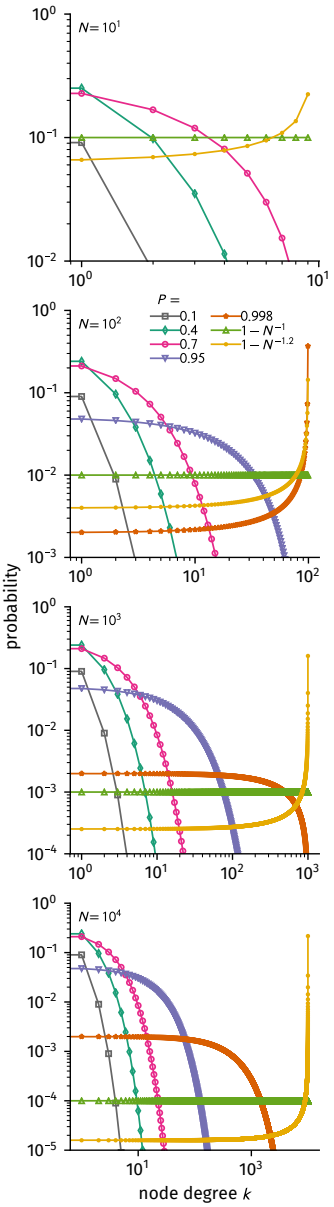


Figure 7.8: Flockwork equilibrium degree distributions as computed from Eq. (7.6) for increasing reconnection probability P and number of nodes N.

This means that ignoring changes in group size of both nodes, an edge attached to two nodes of degree k decays with rate

$$\omega_k^- = 2\gamma \left(1 - \frac{kP}{N-1}\right) \tag{7.11}$$

The probability that any random edge drawn from the network is attached to two nodes of degree k is

$$\pi_k = \frac{kp_k}{\langle k \rangle}. \tag{7.12}$$

Hence, the mean decay rate of a single edge $\langle \omega^- \rangle = \sum_{k=0}^{N-1} \omega_k^- \pi_k$ can be approximated as

$$\langle \omega^- \rangle = 2\gamma \left(1 - \frac{P}{N-1} \frac{\langle k^2 \rangle}{\langle k \rangle}\right). \tag{7.13}$$

This result is merely an approximation because while both nodes are connected, their degree may change through group change events. For small-sized groups, the probability that a group size-change event is also an edge decay-event is higher than for larger groups, so for large values of P (higher mean group sizes), Eq. (7.13) should display a less accurate behavior.

The contact life-time distribution may be approximated as

$$p(\tau_C) = \langle \omega^- \rangle \exp(- \langle \omega^- \rangle \tau_C). \tag{7.14}$$

In App. C.3, a numerical test shows that even for large values of $P > 1 - N^{-1}$ (large groups) Eq. (7.13) yields a highly accurate estimation of the true decay rate. However, for those values of P, the estimation Eq. (7.14) underestimates the tail of the true distribution which decays slower, as will be demonstrated in the following subsection.

In order to map a Flockwork model to the parameters of the edge activity model (introduced in Sec. 2.4.2), one may consequently use Eq. (7.13) and the network density which is

$$p = \frac{\langle k \rangle}{N-1}.$$

In the α-β-variant where $\gamma = \alpha + \beta$ and $P = \alpha/(\alpha + \beta)$ this extends to

$$\langle \omega^- \rangle = 2\alpha \left(1 - \frac{\langle k^2 \rangle}{(N-1)\langle k \rangle}\right) + 2\beta. \tag{7.15}$$

Hence, two parameters which fully define an instance of an edge activity model are directly determined by any two parameters which fully define an instance of the Flockwork model. This is will be of relevance to properly compare disease spreading on both models in Ch. 8.

7.1.5 Comparing the Basic Poissonian Model to Real Data

In order to test Eqs. (7.9)-(7.14), single Fockwork simulations have been performed for $N = 25$ nodes and $P \in \{0.1, 0.7, 1 - 1/25\}$. For

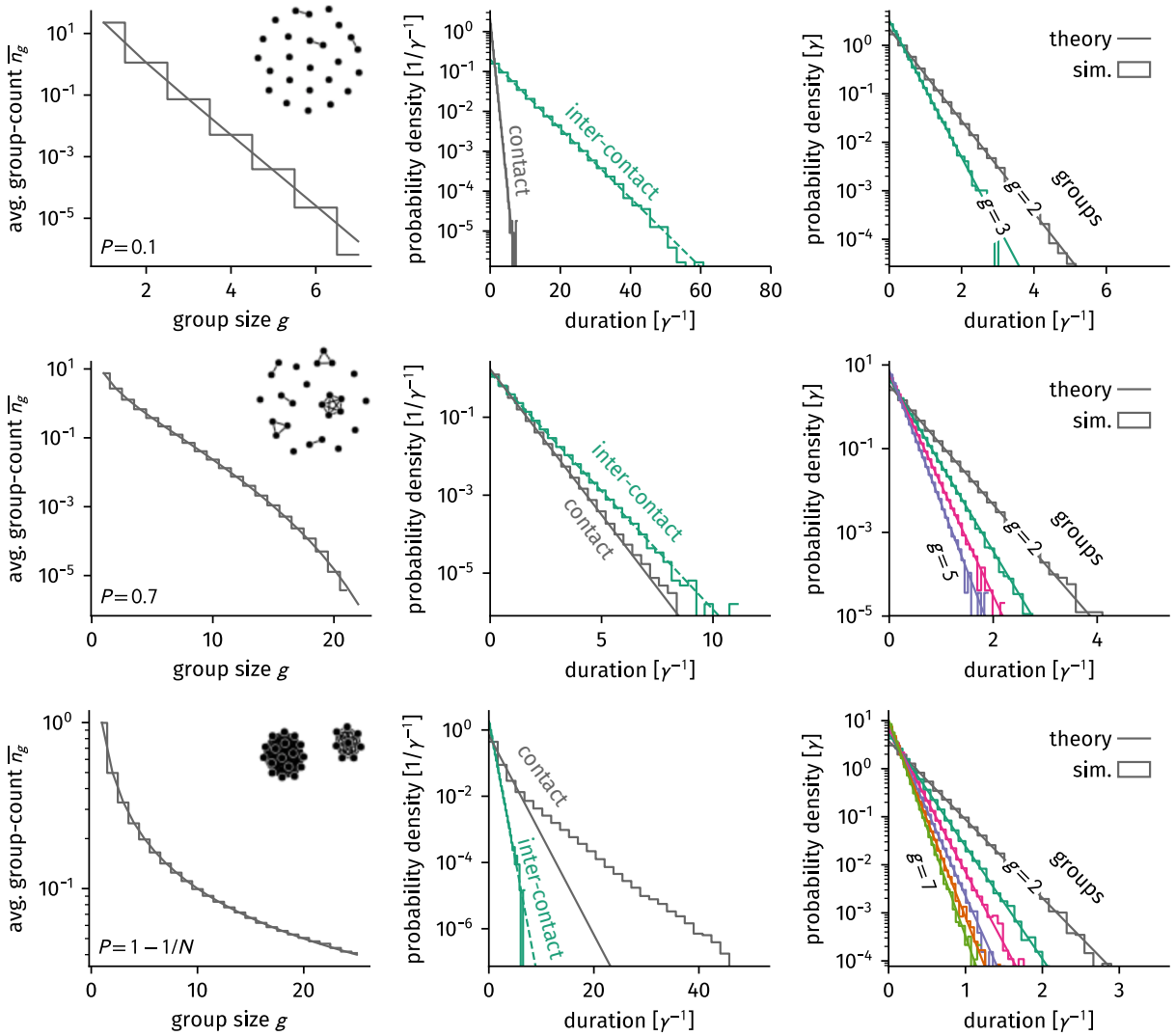


Figure 7.9: Ergodic group-size and life-time distributions of single Flockwork simulations for $N = 25$ and increasing reconnection probability P. Displayed are measurements over a runtime of $T = 3 \times 10^6 (N\gamma)^{-1}$ with $\gamma = 1$. Step functions are histograms and continuous lines represent the theoretical predictions of the group-size distribution Eqs. (7.2), the inter-contact time distribution Eq. (7.9), the approximated contact time distribution Eq. (7.14) and the group life-time distribution Eq. (7.10). Measurements were performed as described in Sec. 2.3.2. The insets in the left column show example snapshots from the corresponding dynamic equilibrium.

each value of P, an equilibrium configuration was sampled using Alg. 1 using the "binomial" and "shuffled group sizes" options. Afterwards, the Flockwork process was simulated for a run time of $T = 10^5 (N\gamma)^{-1}$ to achieve a true equilibrium configuration. Subsequently, the Flockwork process was run until $T = 3 \times 10^6 (N\gamma)^{-1}$ during which the group size distribution, contact and inter-contact time distributions, and group life-time distributions were measured as described in Sec. 2.3.2, with the results shown in Fig. 7.9. For each value of P, the theoretical distributions match with the empirical distributions, with the exception of the contact-duration distribution for large values of P which is predicted correctly for short durations but follows a slower exponential decay in the tail. However, a demonstration in App. C.3 shows that the mean decay rate $\langle \omega^- \rangle$ is still accurately modeled, despite this discrepancy in the tail. Also, the decay in the tail is still of exponential shape, hence not heterogeneous as per the definition established in Sec. 2.1.3.

As discussed in Sec. 2.3.3, real temporal contact networks show

heavy-tailed group-size distributions, as well as heavy-tailed contact time, inter-contact time, and group life-time distributions. For the group life-time distributions, their shape was dependent on the group size, while their cut-off was not. While the Poissonian network model may be tuned to exhibit a power-law group-size distribution, all other behavior is not observed. This is not very surprising as the network model is Poissonian. However, it is well known that Poissonian systems with varying rates can be responsible for heavy tails in inter-event time distributions [47]. Furthermore, real-world contact data shows circadian behavior in their activity. It might therefore be the case that heavy-tailed statistics are a result of circadially varying node activity rates $\alpha(t)$ and $\beta(t)$. This hypothesis will be explored in the following section.

7.2 Real-World Systems as Based on the Flockwork Model

All real-world temporal face-to-face networks introduced in Sec. 2.3.3 show that network activity and number of edges follow a time-varying circadian pattern, which is congruent with expectations obtained by reflecting on everyday life: the majority of humans are far more likely to connect to various other people during the day than they are during the night.[3] It therefore seems naïve to assume that any data measured from real-world interactions might be interpretable as generated from a basic Poissonian Flockwork model. Instead, the following picture will be explored in this section. At any time t of the measurement, the system might be interpreted to be described by two state variables: the active reconnection rate per node $\alpha(t)$ and the average mean degree $k_0(t)$. It will be shown that those quantities can be measured from data, which would allow for an interpretation of a human contact system as being locally equilibrated in time and fully described by the state variable pair (α, k_0), similar to a gas which might be interpreted to go through various equilibrium states in e.g. a Carnot machine [134, p. 604].

[3] The author, spending most of his life in a densely populated urban area, is aware that there are indeed exceptions to this rule, but those will not be of concern here.

7.2.1 Inferring Time-Varying Activity Rates and Network State Variables

Following this argumentation, it is reasonable to assume that the activity rates defined in Sec. 7.1.1 are not constant but time-varying such that $\alpha \to \alpha(t)$ and $\beta \to \beta(t)$.

As introduced in Sec. 7.1.1, the rate α quantifies the average number of active reconnection events a node undergoes per unit time. The rate β is the average number of disconnection events per unit time.

Further, the average mean degree k_0 is the temporal average of the mean degree $\langle k(t) \rangle = \sum_{u=1}^{N} k_u(t)$ as

$$k_0 = \frac{1}{T} \int_0^T \langle k(t) \rangle \, dt.$$

Now, adding a time dependency to k_0, its interpretation is as follows: for any functions $\alpha(t)$ and $\beta(t)$ given a large number of realizations, $k_0(t)$ is the average of $\langle k(t)\rangle$ over all those realizations at time t. The expected number of edges in the system is then given as $e_0 = k_0/(2N)$. Suppose a temporal face-to-face contact network was generated based on a Flockwork process with time-varying rates $\alpha(t)$ and $\beta(t)$. Then, the expected number $e_0(t)$ of edges present in the network will change approximately as

$$\partial_t e_0(t) = \underbrace{-\alpha N k_0(t) - \beta N k_0(t)}_{\text{edges leaving}} + \underbrace{\alpha N(1+k_0(t))}_{\text{edges being established}} . \tag{7.16}$$

In the following, $\langle e_{\text{out}}(t, t+\Delta t)\rangle$ denotes the expected total number of edges being cut in the time interval $(t, t+\Delta t]$ and $\langle e_{\text{in}}(t, t+\Delta t)\rangle$ will be the expected total number of edges being established in the time interval $(t, t+\Delta t]$. Then, according to Eq. (7.16), for a single realization of a Flockwork model, one finds

$$\langle e_{\text{out}}(t, t+\Delta t)\rangle = N \int_t^{t+\Delta t} \left[\alpha(t') + \beta(t')\right] \left\langle k(t')\right\rangle dt'$$
$$\langle e_{\text{in}}(t, t+\Delta t)\rangle = N \int_t^{t+\Delta t} \alpha(t') \left[1 + \left\langle k(t')\right\rangle\right] dt'.$$

As a simplification, one may assume that $\alpha(t)$ and $\beta(t)$ change every Δt but will be constant in between. Then,

$$\begin{aligned}\langle e_{\text{out}}(t, t+\Delta t)\rangle &= N\left[\alpha(t, t+\Delta t) + \beta(t, t+\Delta t)\right] \int_t^{t+\Delta t} \left\langle k(t')\right\rangle dt' \\ &= 2[\alpha(t, t+\Delta t) + \beta(t, t+\Delta t)]E(t, t+\Delta t) \qquad (7.17)\\ \langle e_{\text{in}}(t, t+\Delta t)\rangle &= N\alpha(t, t+\Delta t) \int_t^{t+\Delta t} \left[1 + \left\langle k(t')\right\rangle\right] dt' \\ &= \alpha(t, t+\Delta t)\left[N\Delta t + 2E(t, t+\Delta t)\right], \qquad (7.18)\end{aligned}$$

where

$$E(t, \Delta t) = \int_t^{t+\Delta t} e(t')dt' \tag{7.19}$$

with $e(t)$ being the number of edges present in this realization of the network at time t. Note that E has the dimension of time.[4] From any temporal network the observables e_{in} and e_{out} can be measured by counting the number of edge-creation (edge-deletion) events of the whole network occuring within the time-interval $(t, t+\Delta t]$ as per Eqs. (2.13). The observable E can be measured directly by computing the integral Eq. (7.19).

In order to simply the inference procedure, time will be cut into equally-sized intervals of length Δt such that each rate $\lambda \in \{\alpha, \beta\}$ can be expressed as

$$\lambda(t) = \lambda_n, \text{ for } n\Delta t < t \leq (n+1)\Delta t$$

[4] Even though the number $e(t)$ of edges is changing as a step function, integrals are being used here because the method can be applied to temporal networks of continuous time were the inter-event time is not constant. Thus, the integral simplifies the notation.

with any non-negative integer n. Solving Eq. (7.17) and Eq. (7.18) for α and β, the rates are consequently given as

$$\alpha_n = \frac{e_{\text{in}}(n\Delta t, (n+1)\Delta t)}{N\Delta t + 2E(n\Delta t, (n+1)\Delta t)}$$
$$\beta_n = \frac{e_{\text{out}}(n\Delta t, (n+1)\Delta t)}{2E(n\Delta t, (n+1)\Delta t)} - \alpha_n.$$

Note that from the definition above it might occasionally happen that $\beta_n < 0$, namely when there is no edges being cut but edges being built. In this case one sets $\beta_n = 0$. This is an artifact from actual temporal networks not exactly following the Flockwork node behaviors: In real networks it might happen that only isolated nodes build edges, i.e. choose a group to join. However, in the Flockwork model, the group joining behavior is not limited to nodes in isolation.

Furthermore, for networks where for some times t the number of edges approaches $e(t) = 0$, the calculated rates may yield both $\alpha = 0$ and $\beta = 0$. This can be problematic in simulations based on these rates: The Flockwork will not equilibrate to approach an empty network fast enough and then still have groups whereas the original network is empty. To counter this effect, the disconnection rate β will be artifically increased in order to ensure that every node in the network becomes isolated. The following rule is introduced:

If in both, a time bin $(n\Delta t, (n+1)\Delta t]$ and its following time bin $((n+1)\Delta t, (n+2)\Delta t]$ there is no present edges ($e = 0$), and both rates evaluate to $\alpha_n = \beta_n = 0$, and the previous time bin $((n-1)\Delta t, n\Delta t]$ was not used for emptying, redefine the disconnection rate to be

$$\beta_n = \beta^{\text{emptying}} = \frac{\log N}{\Delta t}.$$

Picking any node at random, it will on average take $N \log N$ picks to reach any node at least once (this is equal to the mean cover time of a complete network, see Sec. 6.1.2). The total disconnection rate $N\beta_n^{\text{emptying}}$ will therefore ensure that on average every node is disconnected during the time interval Δt and will stay isolated afterwards.

Assuming that a temporal network follows the Flockwork model, one may choose two parameters to describe the current state of the temporal network. Since the network has both a structural and temporal component, it makes sense to let this state space be composed of a structural dimension and a temporal dimension. Based on the description above, the parameters α and k_0 are chosen to reflect these dimensions for the following reasons.

The rate β is an unfavorable choice since it is occasionally artificially set and just reflects nodes going to isolation. In contrast, the rate α is always measurable and reflects a physical measure for mixing in the system: it represents the average number of times any node will choose a new environment per unit time. Structurally, one might also use the reconnection probability $P(t)$ as a defining parameter, instead of the expected mean degree k_0. However, considering

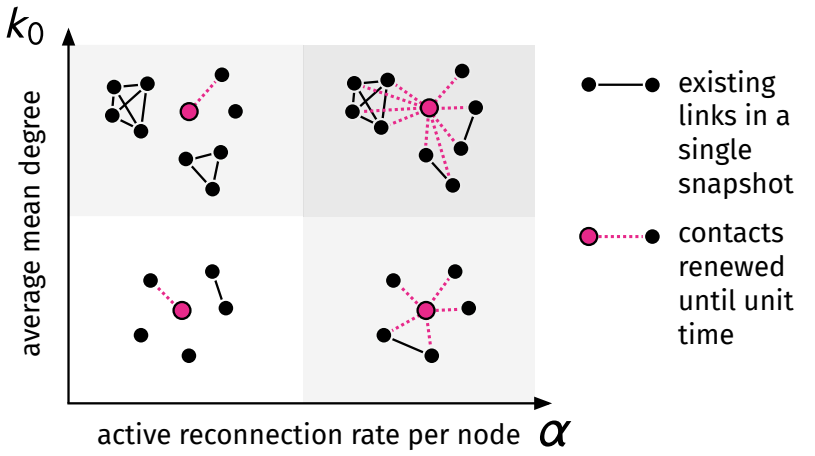


Figure 7.10: Example network behavior for different positions in the Flockwork state space. Solid black lines represent existing edges in a single snapshot, dashed pink lines represent new contacts a single focal node will built (and eventually cut) during an observation of unit time length. Small reconnection rate α implies that a node changes its neighborhood less often per unit time than for a large reconnection rate.

the node-centric view of the temporal dimension, it is intuitive to instead choose a node-centric structural dimension, too: k_0 reflects the average number of connections a node has to other nodes. Therefore it describes how strong a node mixes within the system at a current time. Furthermore, k_0 evolves over time following Eq. (7.16). Thus, assuming equilibrium states are quickly approached, the reconnection probability P can be determined as $P(t) \approx k_0(t)/(1+k_0(t))$ based on Eq. (7.8). One might also wonder why the actual mean degree $\langle k(t) \rangle$ of a configuration at time t should not be chosen as a state variable. This choice is not favorable since this quantity is subject to noise in any Flockwork equilibrium state, while the expected mean degree k_0 is not.

Note that for any position in the state space (α, k_0), typical behavior of the temporal network can be described. Please refer to the qualitative regions displayed in Fig. 7.10. For small values of α, a single node will establish a small amount of new contacts within unit time. These contact events will not produce many edges when k_0 is small, so the network will be very sparse with a large amount of isolated nodes and within unit time the network structure will not change too much. In contrast, for small α and large k_0, larger groups will form, but the network will still appear relatively static when observing for unit time. This changes for large values of α and small values of k_0. Here, the network will still consist of many nodes in isolation, however the network structure will change very much when observing for unit time. Within this time, a single node will establish many new connections but will break them off again within a short time such that k_0 is small at any time. For large values of both α and k_0 large groups exist, but the network structure will change drastically while observing for unit time. A real-world system might be interpreted as following a trajectory in this state space.

In order to obtain these state variables from data, an inference procedure is proposed as follows.

1. Choose a time interval size Δt and cut the observation time of a data set in intervals of length Δt, where each interval is denoted

by a number $n \in \mathbb{N}$ with $n < T/\Delta t$. The last interval size $\Delta t_{\text{final}} \leq \Delta t$ may differ to ensure that the whole observation time $t_{\max}$ is covered.

2. Measure the number of edges e_{in} being created and the number of edges e_{out} being deleted during each time interval $(n\Delta t, (n+1)\Delta t]$ or $(n\Delta t, t_{\max}]$. For each time interval, integrate the number $e(t)$ of edges present during this interval using Eq. (7.19) to obtain E.

3. For each time interval, calculate the rates using Eqs. (7.20) while ensuring that $\beta_n \geq 0$ and that the empty network may be reached by setting $\beta_n = \beta^{\text{emptying}}$ if applicable as described above. Make sure to replace Δt with Δt_{final} for the last interval.

4. Evaluate the function $k_0(t)$ by integrating Eq. (7.16) using $k_0(0) = 2e(0)/N$ where $e(0)$ is the number of initial edges in the analyzed data set.

Unfortunately, it is unclear how Δt has to be chosen. However, one may assume that this procedure infers the correct rates from a simulated instance of a Flockwork model which was generated using predetermined rates α and β changing their value as step functions every Δt_{orig}. Such a test is done in App. C.4 to reveal that rates are correctly inferred if $\Delta t_{\text{inference}} \leq (1/2)\Delta t_{\text{orig}}$ However, for real data, there is no reason for the time-varying rate functions to be step functions changing every Δt_{orig}, so there is no *single* Δt_{orig} to use as an orientation point to choose the right $\Delta t_{\text{inference}}$. A thorough and quantitive method to find the optimal inference parameter $\Delta t_{\text{inference}}$ will thus have to be the subject of future research. Rather, one may learn from the analysis in App. C.4 that it is good practice to steadily increase the inference parameter while observing how the trajectories behave. At small values of $\Delta t_{\text{inference}}$ the trajectories will be subject to strong fluctuations. One may then steadily increase $\Delta t_{\text{inference}}$ and choose a value which reduces the noise while striking features of the noisy trajectories are still present. Furthermore, it is appropriate to compare the measured curve of $\langle k(t) \rangle$ to the inferred function $k_0(t)$ and choose a Δt which satisfyingly reproduces important features of the original curve while smoothing out heavy fluctuations. This method is used in the following to obtain the state-space trajectories of all data sets introduced in Sec. 2.3.3.

7.2.2 *Application to Real-World Face-to-Face Systems*

In this section, the DTU, HT09, and HS13 data sets as introduced in Sec. 2.3.3 are investigated for their state-space trajectory in the (α, k_0)-plane. The analysis is based on the assumption that those data sets can be interpreted to follow the logics of a time-varying Flockwork model.

All three of the data sets are similar in the sense that they contain contact data which has been binned to discrete time intervals Δt, meaning that if an edge was active at any time during the time-interval $[n\Delta t, (n+1)\Delta t)$ it will be considered to have been active

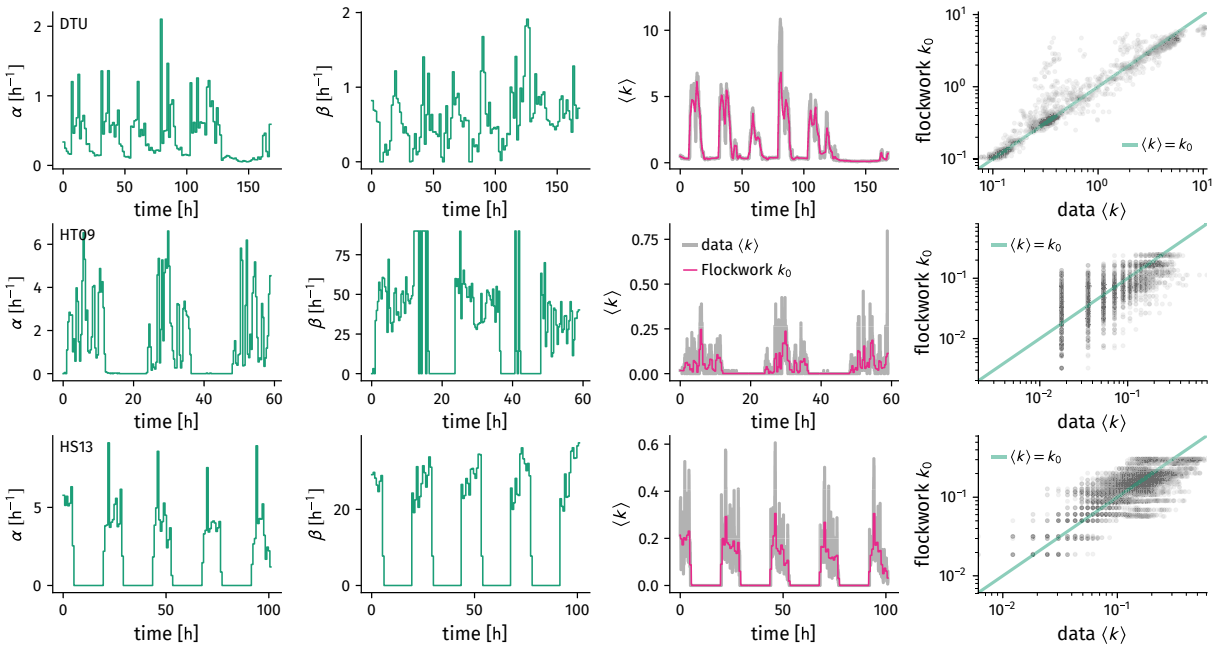


Figure 7.11: Inferred time-varying rates and expected mean degree of the real-world data sets **(top row)** DTU with $\Delta t^{\text{DTU}}_{\text{inference}} = 80\,\text{min}$, **(middle row)** HT09 with $\Delta t^{\text{HT09}}_{\text{inference}} = 20\,\text{min}$, and **(bottom row)** HS13 with $\Delta t^{\text{HS13}}_{\text{inference}} = 45\,\text{min}$. The activity rates $\alpha(t)$ and $\beta(t)$ as well as the expected mean degree $k_0(t)$ show the expected circadian behavior.

through-out the whole time interval.[5] For the DTU data, this value is $\Delta t = 5\,\text{min}$ [79], while for both *SocioPatterns* data sets (HT09 and HS13), the interval length is $\Delta t = 20\,\text{s}$ [40]. One other difference between the data sets is given by the way interactions are recorded. In the *SocioPatterns* experiments, face-to-face contacts actually reflect face-to-face interactions within a radius of $r = 1 - 1.5\,\text{m}$ because the signals which are transferred between chips recording the interaction are shielded by human bodies [40]. In contrast, data from the DTU experiment reflects contacts between all individuals which are within a radius of $r \approx 1.5\,\text{m}$ [79]. Both these differences will be of importance in the following analysis.

For all three data sets, the method introduced in Sec. 7.2.1 was applied for increasing $\Delta t_{\text{inference}}$ where the parameter was chosen as a multiple of the original binning time-interval length. All inferred rate functions as well as the expected mean degree function for varying $\Delta t_{\text{inference}}$ are shown in Fig. F.3-F.11 of App. F. Comparing $\alpha(t)$, $\beta(t)$ and $k_0(t)$ for varying $\Delta t_{\text{inference}}$ shows that the position and shape of the state-space trajectory does not drastically vary with the choice of the inference parameter over a wide range, and as such the exact choice of $\Delta t_{\text{inference}}$ does not seem to be too important. Hence, based on the criteria established in the previous sections (reduced noise, but stable shape of trajectory), the following choices were made: $\Delta t^{\text{DTU}}_{\text{inference}} = 80\,\text{min}$, $\Delta t^{\text{HT09}}_{\text{inference}} = 20\,\text{min}$, and $\Delta t^{\text{HS13}}_{\text{inference}} = 45\,\text{min}$, which yield the step functions shown in Fig. 7.11. For all data sets, the evaluated expected mean degree $k_0(t)$ follows the data while not overfitting to noise (third column and fourth column of Fig. 7.11). Comparing the panels of the first column, the activity rates differ between the DTU data and the SocioPattern sets, showing active reconnection rate values of $0 < \alpha^{\text{SocPat}} \lesssim 6\,\text{h}^{-1}$ per node and $0 <$

[5] with n being a non-negative integer

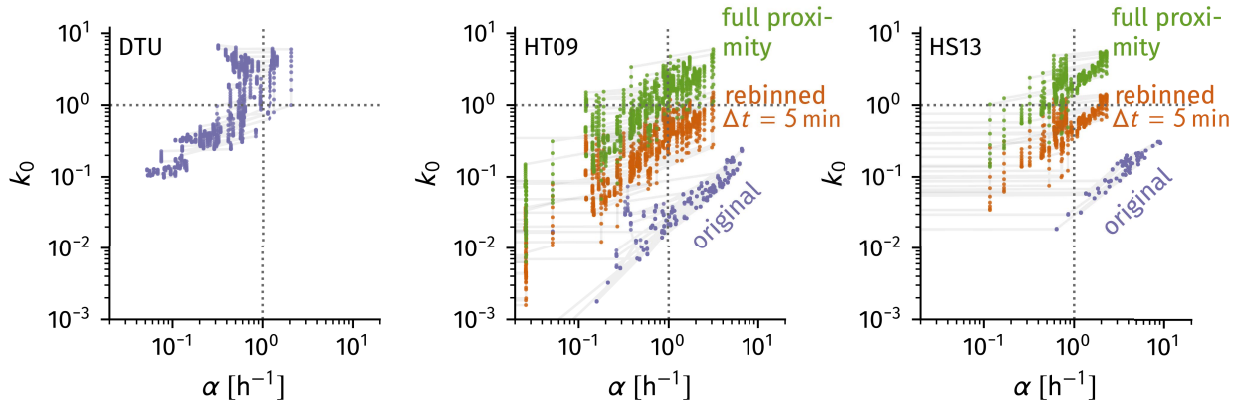


Figure 7.12: Inferred state-space trajectories of the real-world data sets DTU, HT09, and HS13, with equal axis limits to ease comparisons. The trajectories (violet, 'original') show that the networks are changing between slow/sparse states and fast/denser states but are not in slow/dense states or fast/sparse states. This corresponds to the observation that edge count and edge change counts vary in the same rhythm (see panel (d) of Figs. 2.25 and 2.27). While the inferred trajectories of HT09 and HS13 roughly correspond, the DTU trajectory approaches slower and denser states. This discrepance between the DTU and the SocioPattern experiments can be explained by the experimental differences: The DTU data is binned to intervals of length $\Delta t = 5\,\mathrm{min}$. Therefore rebinning the SocioPattern curves to this value increases correspondence (violet to orange states) of the inferred trajectories. Correcting for the fact that the DTU data captures more edges due to measuring proximity instead of face-to-face contacts further increases the shape of the trajectories between the experiments (orange to green states). Note that after both corrections, all three temporal networks qualitatively follow a similar pattern of approached states.

$\alpha^{\mathrm{DTU}} \lesssim 1.5\,\mathrm{h}^{-1}$ per node during the day and values of $\alpha \ll 1\,\mathrm{h}^{-1}$ per node at night, with mean rates of $\overline{\alpha^{\mathrm{DTU}}} = 0.4\,\mathrm{h}^{-1}$, $\overline{\alpha^{\mathrm{HT09}}} = 1.4\,\mathrm{h}^{-1}$, and $\overline{\alpha^{\mathrm{HT09}}} = 1.7\,\mathrm{h}^{-1}$ averaged over the whole measurement. Comparing the panels of the second column, the active disconnection rate β has average values of $\overline{\beta^{\mathrm{DTU}}} = 0.6\,\mathrm{h}^{-1}$, $\overline{\beta^{\mathrm{HT09}}} = 31\,\mathrm{h}^{-1}$, and $\overline{\beta^{\mathrm{HT09}}} = 12.5\,\mathrm{h}^{-1}$ averaged over the whole measurement.

These inferred rates yield state-space trajectories which are qualitatively consistent between the SocioPatterns data sets HT09 and HS13 showing behavior in which the system changes from both low values of α and k_0 to higher values of both in a close to linear fashion (see Fig. 7.12, violet curves, 'original'). The DTU trajectory explores regions where the active reconnection rate is not as large but higher values of the expected mean degree are reached, as compared to the other experiments. This discrepancy can be explained by the two striking methodological differences of the experiments: The SocioPattern data is binned in time bins of $\Delta t = 20\,\mathrm{s}$ and only records face-to-face contacts while the DTU experiment bins data in intervals of length $\Delta t = 5\,\mathrm{min}$ and defines contacts via proximity. Binning data with larger time intervals has the effect of building edges in times were no edges have been observed (see Fig. 2.24), which supposedly increases the expected mean degree. Furthermore, contacts that switch on and off within a binning interval will not contribute to the total amount of edges leaving or being created in this time and as such the rate α is hypothesized to decrease by rebinning data in larger time intervals.

One observes that rebinning the SocioPatterns data set to bins of length $\Delta t = 5\,\mathrm{min}$ and subsequently re-analyzing this new data to obtain $k_0(t)$ and $\alpha(t)$ indeed yields decreased rates and increased mean degrees; the trajectories lie closer to the DTU data (violet curves to orange curves in Fig. 7.12), however, not reaching the high values of the mean degree in the DTU data set. This can be explained by considering the difference between face-to-face data and proximity data. In a disc around a focal individual, face-to-face contacts can only be produced to individuals within roughly one half of this disc, due to signals of the recording RFID chips being absorbed by hu-

man bodies. On the other hand, proximity measurements will count connections to all individuals within the disc. Furthermore, in RFID measurements only individuals whose half disc is turned towards the focal individual will be recorded as contacts. Further assuming homogeneous distribution of people, one can estimate that the expected degree of an average node will be $(1/2) \times (1/2) = 1/4$ of a node in proximity data of similar range. This hypothesis was tested by further multiplying the trajectory of the rebinned SocioPattern data with a factor 4 which finally yields trajectories which are of similar shape as the DTU trajectory (see green curves in Fig. 7.12). This implies that all three systems might follow similar trajectories but the trajectory of the DTU data is altered by the different measurement method. One remarkable feature of the SocioPatterns trajectories is that neither regions where the network is dense but changes slowly, nor regions where the network is sparse but changes fast are explored. It seems that a fast change in network structure is always connected to a more densely connected network and a slowly changing network is associated with a sparse structure which is in agreement with the observation that the number of edges in a network follows the same circadian rhythm as the number of edges being established and being cut as can be seen in the corresponding panel d of Figs. 2.25 and 2.27. This will be of importance when analyzing the spread of diseases for different time scales and configurational settings in Ch. 8.

After analyzing the supposed trajectory a Flockwork would take if it followed activity rates consistent with real world data sets, it is of course of interest whether actual Flockwork simulations follow the structural and temporal statistics seen in real data. To this end, an approximate varying-rate Flockwork simulation algorithm described in Sec. C.4 was run[6] using the inferred rates $\alpha(t)$ and $\beta(t)$ for the DTU, HT09, and HS13 data considering the respective number of nodes $N_{\mathrm{DTU}} = 412$, $N_{\mathrm{HT09}} = 113$, and $N_{\mathrm{HS13}} = 327$. To reflect the way in which the original data is measured, the simulated temporal networks were binned in time bins of length $\Delta t = 5\,\mathrm{min}$ for the DTU surrogates and $\Delta t = 20\,\mathrm{s}$ for the SocioPatterns simulations and compared to the statistics obtained from the non-binned original data. Results for single representative measurements can be seen in Figs. 7.13- 7.15. Panels a of the referenced figures are examined first, where single snapshots of both the original data as well as the binned Flockwork surrogate are shown for two different time points. In general, Flockworks following the varying rates show similar behavior as the original temporal networks, nodes are organized in densely connected groups whose size distribution varies at different times. Furthermore, as can be seen in panel b of the respective figures, the general shape of the number of newly observed contacts as well as the density of edges corresponds between original data and binned Flockwork surrogates. However, more contacts are explored in the Flockwork surrogates which are not seen in the original data, roughly twice as many. Additionally, the real data clearly

[6] The approximate part of this algorithm is that when a rate change happens, the simulation is restarted with the new rate, ignoring all time that has passed before the rate change happened. Using the inference time interval length defined above, this implies that for the DTU data, there will be 126 changes and thus 126 slightly too long inter-event times of a total number of events of approximately $(0.4 + 0.6)\,\mathrm{h}^{-1} \times 168\,\mathrm{h} \times 412 \approx 69{,}000$, For the HT09 data set, there is 177 of those slightly skewed inter-event times of 216,000 events and for the HT13 this evaluates to 135 skewed times compared to 270,000 events. Hence, the error made will be negligible for all inferred rates.

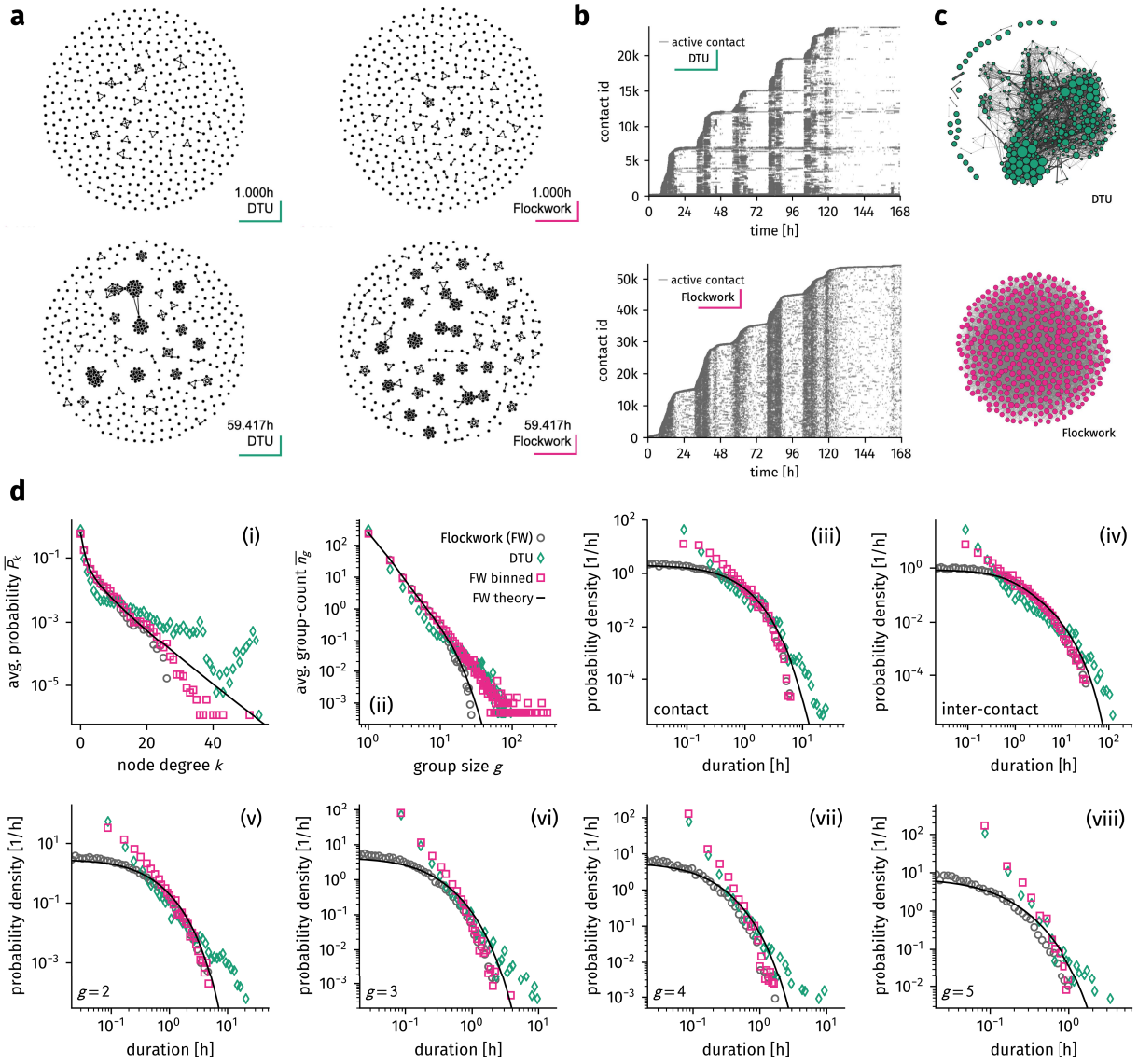


Figure 7.13: Structural and statistical comparison between the DTU data set and a corresponding time-varying Flockwork surrogate. **(a)** Snapshots of the original data and a corresponding binned Flockwork simulation at two different times. **(b)** Edge activity plot. The surrogate explores twice as many edges as the original system and does not produce contacts which are broken and immediately rebuilt. However, the general shape of how those new contacts are explored corresponds as well as the varied network density. **(c)** The aggregated surrogate network is homogeneous and not 'complex' as its original counterpart. **(d)** Temporal distributions of model simulations do not generate the heavy tails observed in real data, however, binning the simulation data produces a heavy tail in group-size distribution and increases correspondence in the degree distribution. Binning also increases the probability of medium-length (inter-)contact life-times. Theoretical predictions Eqs. (7.21)-(7.25) (straight lines) and non-binned data (grey circles) are approximately corresponding.

shows that certain pairs of nodes are either in long contact or tend to break shortly just to be rebuilt almost immediately. This behavior is not replicated by the Flockwork model, in which nodes choose new neighbors uniform at random from the set of all other nodes and as such it is highly unlikely that a recently broken contact is rebuilt in a short amount of time. Another effect of nodes choosing neighbors uniform at random is that each link is equally likely to have been formed in Flockwork surrogates, which yields a homogeneous aggregated network as opposed to the complex ones built from the original data (see panel c of the respective Figs. 7.13-7.15). Here, "homogeneous" refers to the fact that after enough time has passed, each node is connected to every other node with the same weight in the aggregated network. While the model does not replicate the data in this regard the simple limit case of a homogeneous aggregated

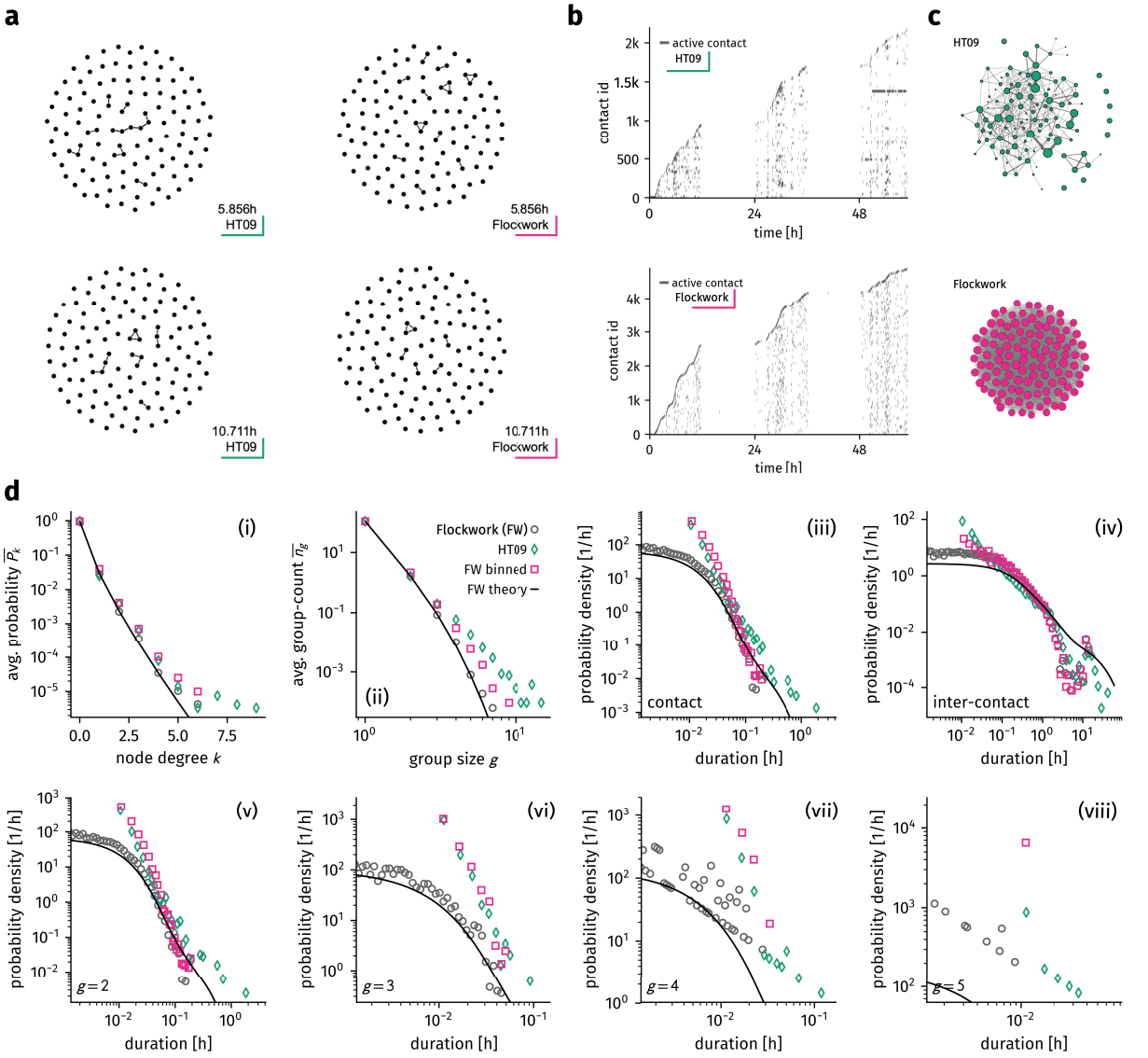


Figure 7.14: Structural and statistical comparison between the HT09 data set and a corresponding time-varying Flockwork surrogate. **(a)** Snapshots of the original data and a corresponding binned Flockwork simulation at two different times. **(b)** Edge activity plot. The surrogate explores twice as many edges as the original system and does not produce contacts which are broken and immediately rebuilt. However, the general shape of how those new contacts are explored corresponds as well as the varied network density. **(c)** The aggregated surrogate network is homogeneous and not 'complex' as its original counterpart. **(d)** Temporal distributions of model simulations do not generate the heavy tails observed in real data, however, binning the simulation data produces a heavy tail in group-size distribution and increases correspondence in the degree distribution. Binning also increases the probability of medium-length (inter-)contact life-times. Theoretical predictions Eqs. (7.21)-(7.25) (straight lines) and non-binned data (grey circles) are approximately corresponding.

network will make it easier to analyze dynamics in aggregated static limit cases on those surrogates.

Subsequently, the statistical properties of structural and temporal observables are analyzed, i.e. the average degree distribution, average group-size distribution, contact duration distribution, inter-contact duration distribution and group life-time distributions of groups of varying size (see panel d of Figs. 7.13-7.15). First, note that these distributions are well-defined for constant-rate Flockwork equilibrium configurations as derived in Sec. 7.1. In the varying-rate Flockwork model, the inferred expected mean degree $k_0(t)$ defines the equilibrium state in a structural sense, which can be approximately mapped to $P(t) = k_0(t)/[k_0(t)+1]$. The temporal state is defined by the inferred rate $\alpha(t)$, such that the effective total event rate will be approximated as $\gamma(t) = \alpha(t)/P(t)$ (which will be equated to $\gamma(t) = 0$ if $k_0(t) = P(t) = 0$). Hence, assuming that the Flock-

work follows a trajectory through the state space, all structural and temporal property distributions should be approximately given by a temporal average over these equilibrium states with $\gamma(t)$ and $P(t)$. For the structural distributions of degree and group sizes, one may use Eq. (7.2) and Eq. (7.6) to find

$$\overline{n_g} = \frac{1}{T}\int_0^T dt\, n_g[P(t)] \tag{7.21}$$

$$\overline{P_k} = \frac{1}{T}\int_0^T dt\, P_k[P(t)]. \tag{7.22}$$

Using Eq. (7.9), the average inter-contact time distribution might be estimated as

$$\overline{p(\tau_{IC})} = \frac{1}{T}\int_0^T dt\, 2\alpha(t)\exp\left[-2\alpha(t)\tau_{IC}\right], \tag{7.23}$$

while the approximated contact-time distribution defined by Eqs. (7.13) and (7.14) evaluates to

$$\overline{p(\tau_C)} = \frac{1}{T}\int_0^T dt\, 2\gamma(t)\left(1-\frac{P(t)}{N-1}\frac{\langle k^2\rangle[P(t)]}{\langle k\rangle[P(t)]}\right)\times \exp\left[-2\gamma(t)\left(1-\frac{P(t)}{N-1}\frac{\langle k^2\rangle[P(t)]}{\langle k\rangle[P(t)]}\right)\tau_C\right]. \tag{7.24}$$

Note that here, $\langle\cdot\rangle[P(t)]$ denotes the configurational equilibrium expectation using $P_k[P(t)]$ as per Eq. (7.6). Finally, one may approximate the group life-time distribution as

$$\overline{p(\tau_g)} = \frac{1}{T}\int_0^T dt\, \left(g\gamma(t)(1-P(t)) + 2\alpha(t)g\frac{N-g}{N-1}\right)\times \exp\left[-\tau_g\left(g\gamma(t)(1-P(t)) + 2\alpha(t)g\frac{N-g}{N-1}\right)\right]. \tag{7.25}$$

When the time of measurement T is sufficiently large and the process truly follows a trajectory through equilibrium states, these temporal averages correspond to averages over rate distributions $p(\lambda)$ as

$$\overline{p(\tau)} = \langle p(\tau)\rangle = \int_0^\infty d\lambda\, p(\lambda)\lambda\exp(-\lambda\tau).$$

However, using the inferred rate functions above, several rates will be $\alpha = \beta = 0$ for a substantial amount of time and thus contribute $p(\tau) = 0$ to the integral for finite times. Intuitively though, the inter-event time between two events will be shifted to the length of the period in which the event rates are zero and thus finite probability for

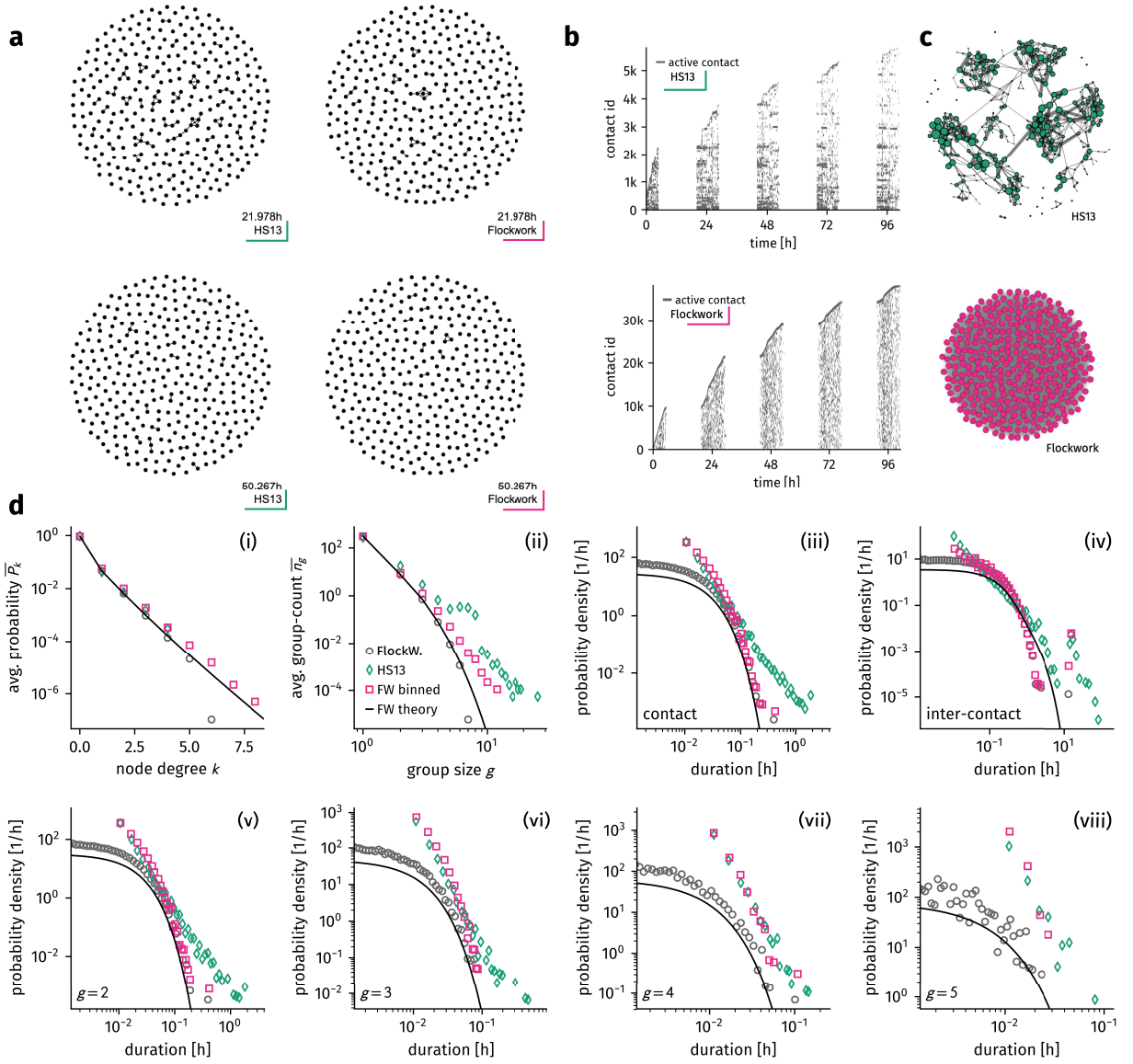


Figure 7.15: Structural and statistical comparison between the HS13 data set and a corresponding time-varying Flockwork surrogate. **(a)** Snapshots of the original data and a corresponding binned Flockwork simulation at two different times. **(b)** Edge activity plot. The surrogate explores twice as many edges as the original system and does not produce contacts which are broken and immediately rebuilt. However, the general shape of how those new contacts are explored corresponds as well as the varied network density. **(c)** The aggregated surrogate network is homogeneous and not 'complex' as its original counterpart. **(d)** Temporal distributions of model simulations do not generate the heavy tails observed in real data, however, binning the simulation data produces a heavy tail in group-size distribution and increases correspondence in the degree distribution. Binning also increases the probability of medium-length (inter-)contact life-times. Theoretical predictions Eqs. (7.21)-(7.25) (straight lines) and non-binned data (grey circles) are approximately corresponding.

longer inter-event times should contribute to the average above. This results in an obvious discrepancy between the theoretical predictions and the unbinned simulation data.

Comparing the structural statistics, one first notices that the predicted degree and group-size distributions Eqs. (7.21) and (7.22) align rather well with the results from simulated data for all data sets as one may see in panels d.i–d.ii of the respective Figs. 7.13, 7.14, and 7.15. The unbinned degree distributions also correspond well with the data degree distributions in the cases of HT09 and HS13, however, they differ in the DTU case which seems to have substantially more nodes of higher degree while still decaying exponentially. Binning the Flockwork surrogates to the time bins of their respective real experiments increases the typical degree-scale, but not drastically. In contrast, the group-size distributions of the unbinned Flockwork data decays faster than a heavy-tailed distributions for all

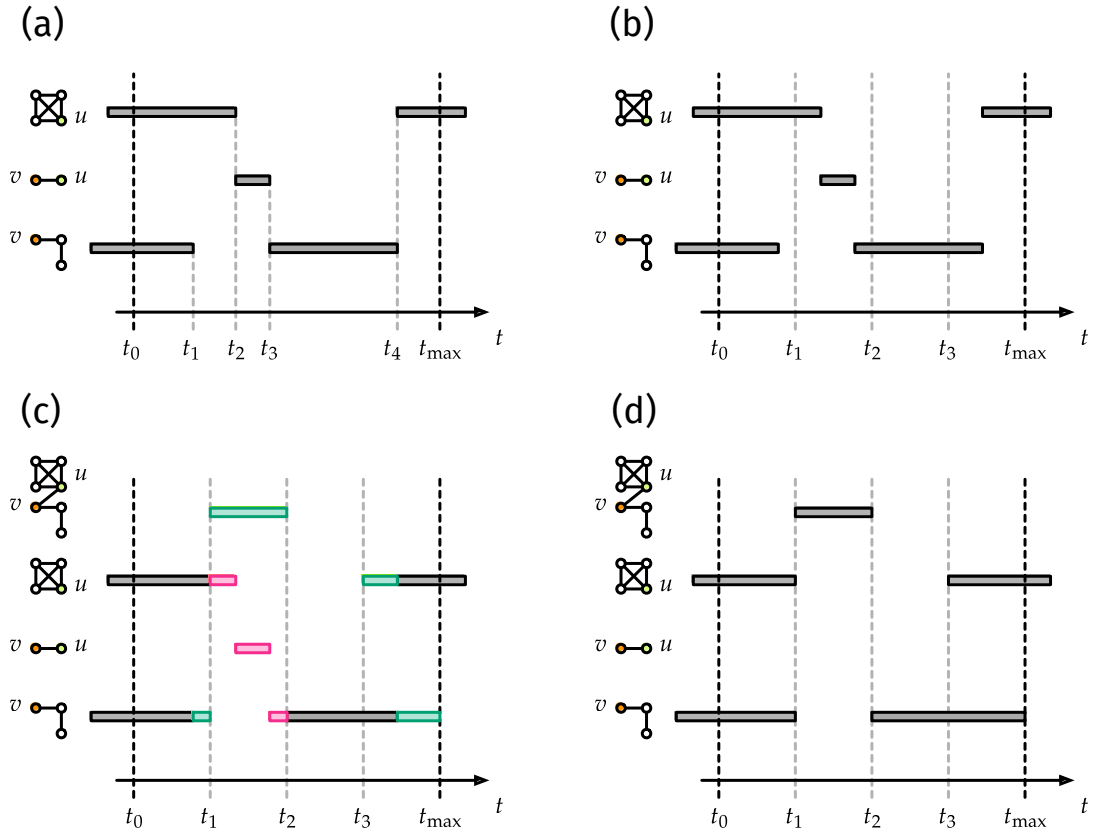


Figure 7.16: The effect of binning on measuring groups. **(a)** An example temporal network recorded in continuous time displayed as existing groups over time. A grey bar means that the node set of this group remained unchanged over the time period it spans. **(b)** Time is discretized. **(c)** Edges are binned according to the binning procedure described in Sec. 2.3.1. This leads to the eradication of smaller groups which are merged to build previously non-existent larger groups. **(d)** The resulting network contains a large group which did not exist before.

data sets, which changes when binning is introduced—binned Flockworks show indeed heavy tails in the group size distributions. Since individuals can change between groups in a short amount of time, binning will lead to groups being measured as connected where they are actually not, thus increasing the amount of larger groups in the system (see Fig. 7.16 and the explanation given in the figure caption) while not strongly changing the shape of the degree distribution. This seems to be true for HT09 and HS13. However, the comparison of degree and group-size distribution of DTU data and binned Flockwork in panel d.i of Fig. 7.13 suggests that binning cannot be the only effect responsible for heavy-tailed group size distributions in this experiment. The larger amount of nodes with higher degrees compared to the binned Flockwork suggests that groups in the original data are indeed larger than in the simulation. It seems that in some single snapshots, the group-size distribution is bimodal with some very large very dense groups and many single nodes. This behavior cannot be replicated with the Flockwork model.

Changing focus to the temporal observables one first notices that the distributions of all unbinned simulations flatten off for short duration, while binning to a fixed time-scale increases the amount of short medium-length durations which leads to distributions which might be mistaken for broad distributions when only analyzing shorter times (see panels d.iii-d.viii of Figs. 7.13-7.15). However, the tails of

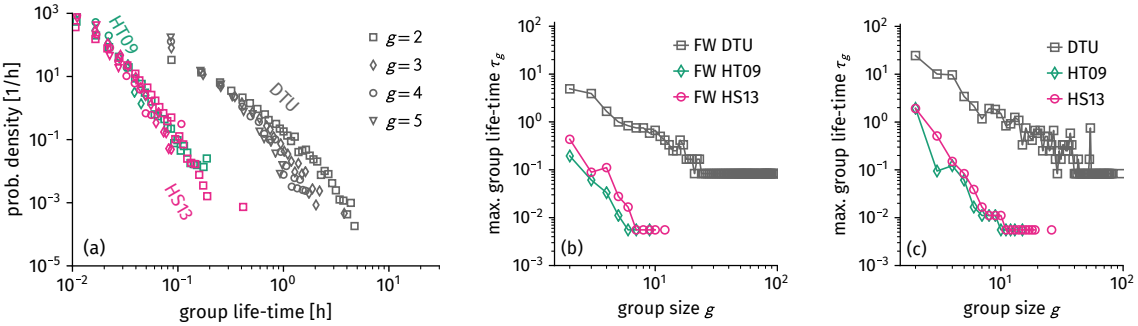


Figure 7.17: Comparison of binned Flockwork group life-time distributions. **(a)** Both the distributions of the SocioPatterns surrogates (HT09 and HS13) correspond to each other and their shape does not vary significantly with group size, replicating behavior observed in real data (cf. Fig. 2.28). For the DTU data, the shape of the group life-time distribution varies with group size. The heavy-tailed decay of the group-size distribution in real data cannot be observed in the surrogates. **(b)** The maximum group life-time does indeed decrease with increasing group size g, similar to the real data (shown in **(c)**).

the binned Flockwork distributions correspond to the tails of the unbinned Flockwork distributions which seem to decay exponentially as opposed to their real-world counterparts which show heavy tails. It thus seems that time-varying rates are not sufficient to explain heavy-tails in the temporal distributions. Qualitative exceptions are observed in the contact and inter-contact time distributions which contain some larger times in HS13 and HT09 (panels d.iii and d.iv of Figs. 7.13-7.15). Considering the inter-contact time, indeed, nodes will not interact at night, thus producing those larger inter-contact times. The group life-time distributions do not show strong heavy tails as the original data does, however, binning can be identified as the cause for these distributions having similar shape in HT09 and HS13 (see Fig. 7.17a). Furthermore, the maximum observed life time of a group decreases with increasing group size (Fig. 7.17b), similar to the original data (Fig. 7.17c). Note that the group life-time distributions of the DTU data indeed change shape with increasing group size even when binned, as opposed to the distributions obtained from the DTU experiment.

The simulations show that some of the properties of real-world face-to-face systems can be reproduced by Flockwork simulations while others cannot. This will be summarized and discussed in detail in the following section.

7.3 *Summary and Discussion*

In this chapter, a simple Poissonian model called 'Flockwork' was introduced, which simulates temporal networks in continuous time. This model entails nodes switching between groups and being isolated by events that happen with a constant active reconnection rate and a constant active disconnection rate. It was shown that regardless of the rate choices as well as the initial condition, a statistical equilibrium state will be reached in which the network consists of fragmented completely connected groups. An algorithm was suggested to generate configurations from the equilibrium ensemble. The group-size distribution of this equilibrium state can be computed, as well as the degree distribution, inter-contact duration distribution, group life-time distributions and an estimated contact du-

ration distribution.

Real-world temporal networks, however, are not found to be in any stable dynamic equilibrium, which was shown for exemplary data sets from the Copenhagen network study (DTU) [39, 79] and the SocioPatterns experiment (HT09 [42] and HS13 [41]). Instead, their activity and density varies over time in a circadian manner and their structural and temporal properties are distributed with heavy tails. Existing theories attribute these heavy tails to memory effects and age-dependent activity rates while the circadian nature of the data is ignored [43, 44]. Heavy-tailed inter-event time distributions, however, maybe a consequence of a system depending on multiple time-scales and as such one might wonder how necessary the consideration of these memory effects is to explain the heterogeneous distributions. In this chapter the hypothesis was explored whether circadian time-varying activity rates are responsible for the heavy tails in the network's structural and temporal properties, ignoring memory effects. To this end, a method to infer time-varying activity rates from real-world data was developed first, which showed that if real-world systems followed Flockwork models with varying rates, they would indeed be expressible to follow trajectories of equilibrium states in a newly found two-dimensional state space. One axis of this state space is given by the global active reconnection rate per node which quantifies the average number of active reconnections a single node undergoes per unit time. The second dimension of this state space is quantified by the expected mean degree of the network which directly translates to a group size distribution and corresponds to the network density. For HT09 and HS13 it was found that their inferred trajectories approximately correspond to each other and vary between states where the network is sparse and changes slowly and states where the network is denser and changes faster, ignoring regions where the network is dense but changes slow or is sparse but changes fast. The DTU data seemed to take a shifted, more erratic trajectory. However, it was shown that the HT09/HS13 systems displayed similar trajectories to the DTU data when rebinned to time bins of 5 min length and considering pure proximity data instead of face-to-face data. This suggests that all the systems analyzed here behave similarly (in a qualitative manner) over the course of a few days. Interpreting real-world human contact networks to be embedded in such a low-dimensional state space offers a new perspective and could help understand how diseases spread in face-to-face systems at different times, when the system is in different equilibrium states.

Simulating Flockworks with these inferred time-varying rates it was further shown that heavy-tailed group-size distributions are produced, which is largely attributed to the experimental practice of binning data. Before data was binned, the group-size distributions decayed faster than heavy-tailed. This insight was previously unexplored since other models of temporal networks were either simulated in discrete time or analyzed unbinned (cf. Sec. 2.4). Hence,

this insight might have implications when recording and analyzing contact data in future experiments. In order to find out whether group sizes actually decay with a heavy-tail, continuous-time measurements could be a key to the answer. In at least one of the experiments, though, the DTU experiment, the right-skewed degree distribution suggests that aggregated groups might indeed be distributed with a heavy-tail even in non-binned data.

Aggregating the surrogate networks showed that no complex structure is observed: node degrees and link weights (representing total amount of time spent together) are homogeneously distributed, as opposed to the aggregated structures from the original data which showed both link and node heterogeneity. This is not surprising as node event rates and connection probabilities are completely homogeneous in the Flockwork model even for time-varying rates.

Analyzing the temporal distributions it was found that binning data skews all observed short durations to the mid-length regions. However, heavy tails are not observed. This lack of heavy tails might be attributed to the node-homogeneity of the Flockwork model. Each node undergoes dis- and reconnection events with the same rate, hence, mixing is relatively fast and contacts are built and decay on a homogeneous time-scale during a certain time of constant activity rates. Hence, introducing node-specific dis- and reconnection rates whose averages follow the inferred rates over time might introduce heavy tails in contact and group-life time distributions. This would be induced by long-lasting contacts of nodes which connect once but are reluctant to disconnect due to their reduced activity rates. Second, the inter-contact durations might be influenced by dropping contact likelihood homogeneity. In the current model, nodes are equally likely to connect to every other node. This is, in fact, a simplified assumption: humans are more likely to establish contacts to people they already interacted with. Reflecting this in simulations respecting link heterogeneity, a node which has a low node-specific connection rate will not establish any contacts and additionally it might not get connected to by another node, presumably leading to higher inter-contact times. One way to achieve link heterogeneity in a system is certainly given by considering memory kernels which control that nodes reconnect only to nodes with which they had already interacted. However, this might be far fetched to simulate data like DTU, HT09, and HS13, where experiments measure systems with an already established social network: recorded people usually already know other nodes and hence contacts are not produced from memory in observed data but rather enforced from an established static social network (memory from pre-experimental data). Thus, another method to reflect this circumstance in a model would be to force nodes to reconnect using a pre-defined static social network, which might further reproduce this static social network in the aggregated limit. However, contacts in real-world data still have the property to be relatively likely to be reestablished in short time once they are broken off [89], which would not happen when enforcing an externally

given social network. Memory seems unavoidable to reflect this network property. However, one key property which was not discussed yet is that human people are operating in space. Hence, people who break off contacts are very likely to reconnect to people which are in higher proximity which might result in the described effect without introducing explicit memory kernels. Indeed, an already popular model generates temporal networks by letting nodes perform a random walk in two-dimensional space and connecting nodes based on proximity [135] which can reproduce heterogeneous group-size distribution and contact durations, however, circadian activity is once again ignored. To the best knowledge of the author, this replication of circadian activity is not entailed in any other temporal network model yet. The model formulation based on node behavioral rates furthermore allows to map its samples to continuous time and thus to real data, which has not been done previously. For the first time, this allows to answer the question why the group-size distribution assumes a heavy-tailed shape: This is an emergent effect of circadian activity in combination with the experimental practice of binning temporally resolved contacts to discrete time. The fragmentation of a system to small, dense groups was further shown to be an emergent properties base on the node behaviors, instead of a postulate as in other models. The Flockwork model is therefore a valuable contribution to the collection of reference models to investigate temporal network properties and their influences on dynamics.

In order to increase the realism of the model, as discussed above, node and link heterogeneity might be investigated in future studies to be responsible to produce observed statistics in real-world data once they are incorporated into the Flockwork model. However, in its current implementation the model is already a valuable *null* model. Here, 'null model' refers to a model which replicates one or more certain aspects of real data while others are ignored such that by comparing its samples to empirical data one can confirm or reject hypotheses. On this account, the constant-rate Flockwork model may be used in the future to study the effect of group-sized topology on the outcome of dynamic processes such as diffusion or disease spreading, restricting the analyses to a homogeneous time-scale to learn about the effect of group structures only. In contrast to previous models (cf. Sec. 2.4), the self-similarity of the Flockwork model decreases as a smooth, homogeneous function in time, allowing for thorough investigations of the influence of time-scales in continuous-time temporal networks.

Further using the time-varying rate formulation of the model, the influence of circadian activity can be studied, without effects imposed by node or link heterogeneity. In the following chapter, one such analysis will be done: the static-rate Flockwork model will be investigated to study the effects of group structures on the spread of a susceptible-infected-susceptible disease dynamic when the model is in different points in the (α, k_0)-plane.

8
The SIS-Model on Temporal Contact Networks

In epidemiology, predicting the number of people affected by an infectious disease is as important as estimating the systemic susceptibility to be vulnerable to such a disease in the first place. For the former, one is interested in evaluating the ratio of infected people if no measures are taken to counter a disease's spread—aiming at e.g. estimating the cost of treatment in a worst-case scenario. The latter is quantified by the epidemic threshold: How much do we have to reduce the number of secondary infections in order to eradicate a disease from the whole system? In the following, the answer to these questions is discussed for temporal network models and real-world data.

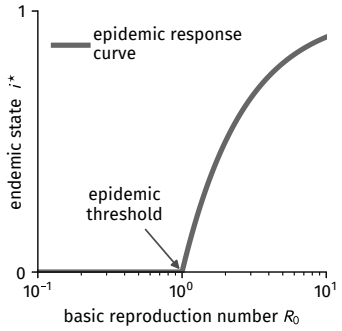


Figure 8.1: A short reminder: the epidemic response curve represents the fraction of infected people in equilibrium, functionally dependent on the basic reproduction number R_0. The value $R_{0,c}$ at which the endemic state $i^\star$ first vanishes is called the epidemic threshold. Above the epidemic threshold, two equilibrium points exist: the unstable disease-free state $i^\star = 0$ and the stable endemic state $i^\star > 0$. Below the epidemic threshold, only the disease-free state exists and is stable.

8.1 Introduction

As introduced in Sec. 3.2, computationally modeling the spread of infectious diseases in human systems proved to be an effective method to perform such estimations. In particular, the susceptible-infected-susceptible (SIS) infectious model provides a simple framework to predict both the severity of a disease given the number of secondary infections as well as the epidemic threshold (both depicted in Fig. 8.1). However, various problems regarding predictions based on this model were discussed, as well.

The classic SIS process is a link-based Poisson process where a single contact between an infected node and a susceptible node decays to a contact between two infected nodes with a constant infection rate. Often, this basic Poisson process is replaced by a reactive process, for which each infected individual tries to infect all or a fraction of its neighbors with a certain probability within a discrete time step of length Δt. In Sec. 3.2 it was shown that this process can be adequately described by a system of discrete-time equations that model the reactive process based on an individual-based Markov-chain approximation. Within this approximation, the covariance of neighboring nodes' infectious states is assumed to vanish. It was further demonstrated that this Markov-chain approximation indeed models the reactive process well on static networks, while it may

fail to coincide with the results of continuous-time SIS simulations: for instance on networks that are relatively sparse and not randomly generated. The assumption of invariant neighboring infection states are not justified in those networks, and therefore predictions based on the Markov-chain approximation should be met with care in general [52].

The critique in [52] assumes a static topology, i.e. that two nodes are connected at all times. However, in Sec. 2.3 and Ch. 7 it was discussed that this is indeed not the case in human face-to-face contact systems: At any point in time, humans are in contact with only a limited number of people, are changing these contacts based on a circadially varying activity, or are isolated.

Recent studies discussing the implications of network temporality on the outcome of SIS dynamics [45, 56, 57, 91] are all based on the discrete-time individual-based Markov-chain approximation and justify this approximation by simulations using the reactive process instead of the basic continuous-time Poisson process.[1] In the studies cited above, the time-scale of a temporal network is defined as the length of a single snapshot.

[1] In [45], it is unclear which simulation algorithm was used but it is implied that it is the reactive process.

With a low recovery rate $\rho \to 0$ an infected node 'sees' many snapshots of a temporal network before recovering. This limit will be referred to as the annealed limit. In this case, the SIS process is equal to one on the corresponding averaged network.

In [45], Speidel *et al.* argue that for increasing $\rho > 0$, i.e. increasing importance of a network's temporal structure as compared to the averaged network, the epidemic threshold decreases, making the system more susceptible for an epidemic. Their finding is supported by simulations[2] to obtain the quasi-stationary (i.e. non-absorbing) endemic state on various real-world and model temporal networks. In the first investigated model, snapshots are constructed randomly and independent from the last snapshot, each consisting of disjoint fully connected groups of homogeneous size and isolated nodes. A single temporal snapshot of their investigated model is therefore similar to a single snapshot of the Flockwork model, which was introduced in the last chapter. Their investigated real-world networks include the HT09 data set introduced in Sec. 2.3.3 and discussed in Sec. 3.2.4. However, all investigated real-world temporal contact data was rebinned to aggregation times much longer than the original resolution. In Sec. 2.3 and Ch. 7 it was already demonstrated that rebinning temporal contact networks significantly changes their structure.

[2] probably sinulating the reactive process, see footnote above

In other studies, Valdano *et al.* showed that the individual-based Markov-chain approximation reproduces simulation results with high accuracy, where the reactive proces was used in discrete time [56, 91]. In these studies, however, real-world networks were also rebinned to windows of much longer time than their actual temporal resolution. Furthermore, the circadian nature of the data was ignored by discarding data recorded during the night [56].

In this chapter, the difference between the individual-based Markov-chain approximation (MCA) and the exact continuous-time simula-

tion on temporal networks is studied. In particular, an exact continuous-time perspective is chosen to simulate the Poissonian SIS process and concurrently the change of the temporal network for the Poissonian Flockwork model and the edge activity model. The results will be compared to the results given by the MCA on these continuous-time networks to reveal that the approaches may yield opposite outcomes for both temporal network models.

In a further analysis, epidemic spreading is studied in detail on the real-world temporal networks DTU, HT09, and HS13, as well as Flockwork surrogates of circadian activity.

8.2 Spreading on Temporal Network Models

In this section, the epidemic response curves of the Poissonian SIS model will be studied for the Poissonian Flockwork model and the edge activity model, both in terms of continuous-time simulations and the individual-based Markov-chain approximation (MCA).

The Flockwork model was introduced in Ch. 7 and models the structure of real-world temporal face-to-face contact networks. Nodes are either connected to small, dense groups, or isolated. They cut their contacts and search a new neighbor with active reconnection rate α. As discussed in Ch. 7, temporal networks may be interpreted to follow a trajectory of equilibrium states in the (α, k_0)-plane over the course of a day. Hence, it is vital to understand how a disease behaves in each of those equilibrium states before analyzing its behavior on the circadially varying networks. In order to obtain an understanding of the Flockwork's structural influence on disease spreading, the results will be compared to simulations on the edge activity model.

The edge activity model was introduced in Sec. 2.4.2 and is the Poissonian contiunation of the Erdős–Rényi (ER) model: At each time, the network consists of a sample of the ER model. Each existing edge is deleted with rate ω^- and each non-existing edge is activated with rate ω^+. The Flockwork model's defining activity rate α can be mapped to the edge activity model's rate ω^- using Eq. (7.15). Both models will be simulated using Gillespie's SSA (see App. D) with (i) a concurrent Gillespie simulation of the SIS dynamic, and (ii) a concurrent integration of the individual-based Markov-chain approximation equations given by Eqs. (3.26) and (3.27).

8.2.1 General Considerations

As stated above, one major task of epidemiology is to predict the impact of a disease for a given topology. To this end, one measures the ratio $i^\star = N^{-1} \sum_{u=1}^{N} \langle i_u^\star \rangle$ denoting the fraction of infected individuals in equilibrium. In this chapter this observable's functional dependence on both varying group-size distribution as well as active reconnection rate α is investigated using Gillespie simulations.

In order to control the network structure of different temporal

network models of different density, the average mean degree

$$k_0 = \frac{1}{T}\int_0^T dt \frac{1}{N}\sum_{i=1}^{N} k_i(t) \tag{8.1}$$

will be used as the *structural control parameter*. It is an obvious choice as it directly controls the number of contacts in a system and is both directly mappable to a Flockwork equilibrium configuration with a certain group-size distribution as well as to the network density in the edge activity model.

The time-scale will be fixed by keeping the active reconnection rate α of a Flockwork model constant for all situations, while the recovery rate per node ρ is varied. The control parameter

$$c_0 = \frac{\alpha}{\rho}$$

then describes the number of active reconnections a node performs before it recovers, on average. In the context of the edge activity model, it refers to the number of times a node changes *all* of its $\langle k \rangle$ neighbors until it recovers. This parameter will hence be called the *contact-renewal number* and controls the temporal change of a single node's local structure in comparison to the time-scale of the disease.

For the edge activity model, the temporal control parameter corresponding to the active reconnection rate α of the Flockwork model is given by the uniform link decay rate ω^- which is determined by Eq. (7.13). In Fig. 8.2c and Fig. 8.3c, the parameter choices for Flockwork and edge activity model, respectively, are illustrated.

Further, one is interested in keeping the infectiousness of a disease comparable while varying the network structure. As discussed in Sec. 3.2, a non-severe disease (i.e. with small constant infection rate η) will die out on a sparse network while it might persist on a densely connected network. Hence, the mean-field basic reproduction number

$$R_0 = k_0 \frac{\eta}{\rho}$$

will be the control parameter of a disease's severity, correcting the infectiousness for the average number of contacts per node.

To sum up, the endemic state described by the ratio $i^\star$ of infected nodes in equilibrium will be studied as a function of the structural control parameter k_0 (the average mean degree), the temporal control parameter c_0 (the contact renewal number) and the disease control parameter R_0 (basic reproduction number):

$$i^\star \equiv i^\star(k_0, c_0, R_0).$$

For $k_0 < 1$ both the edge activity model as well as the Flockwork model produce network configurations composed of small, disconnected groups. When increasing the average mean degree to values $k_0 > 1$ only the Flockwork model produces fragmented networks while the edge activity model possesses a giant component (remember that the ER undergoes a percolation phase transition at $k_0 = 1$

as discussed in Sec. 2.2.2). Hence, SIS simulations will be studied on both models to investigate how the fragmented state of temporal face-to-face contacts influences the disease spreading process with different time-scales.

A few considerations can be done already. If $c_0 \to \infty$, an infected node re-connects to a high number of other nodes before it recovers. Hence, it is expected that independent of the structure of the network, the endemic state will follow the mean-field result

$$i^\star(k_0, c_0 \to \infty, R_0) = 1 - \frac{1}{R_0}.$$

This is due to the fact that in both the Flockwork as well as the edge activity model, each link is equally likely to be activated and to transmit an infection when considering a large observation time, i.e. in the annealed limit. When $c_0 \ll 1$, the network will appear static with respect to the evolving disease, i.e. it might suffice to study the spread on an Erdős–Rényi model of degree k_0 for the edge activity model or a single static configuration from the Flockwork equilibrium group-size distribution, respectively.

In the case of $k_0 \to N - 1$, the network will for both models approach the completely connected network topology and hence

$$i^\star(k_0 \to N - 1, c_0, R_0) = 1 - \frac{1}{R_0}.$$

The second observable to be studied is the epidemic threshold, i.e. the critical basic reproduction number necessary to sustain the disease within the population. Since here, the endemic state $i^\star$ in equilibrium takes a functional form depending on k_0, c_0 and R_0, each of those can be 'critical' when the other two are fixed. In general, the threshold will be given by the surface defined by

$$i^\star(k_0, c_0, R_0)\Big|_{\text{crit}} = 0.$$

In the following, these situations will be investigated numerically using Gillespie simulations on the Flockwork model and the edge activity model.

8.2.2 *Epidemic Response*

Gillespie SIS Simulations were performed to obtain the endemic state $i^\star$ of infected nodes in equilibrium for four values of R_0 on both Flockwork and edge activity model, respectively, while varying both the average mean degree and the contact-renewal number between values of $10^{-2} \lesssim k_0 \lesssim 20$ and $10^{-1} \lesssim c_0 \lesssim 10^2$ for $N = 10^4$ (details concerning the simulation are summarized in App. E.1).

The results for the Flockwork model are displayed in Fig. 8.2. Focusing on panel (a) first, one observes that for all values of R_0, a disease-free state is reached for values of $k_0 \lesssim 1$ and $c_0 \lesssim 1$. This corresponds to situations where the network consists mostly of isolated nodes and a small number of small groups and an infected node recovers before reconnecting to a new neighborhood. Therefore, this is

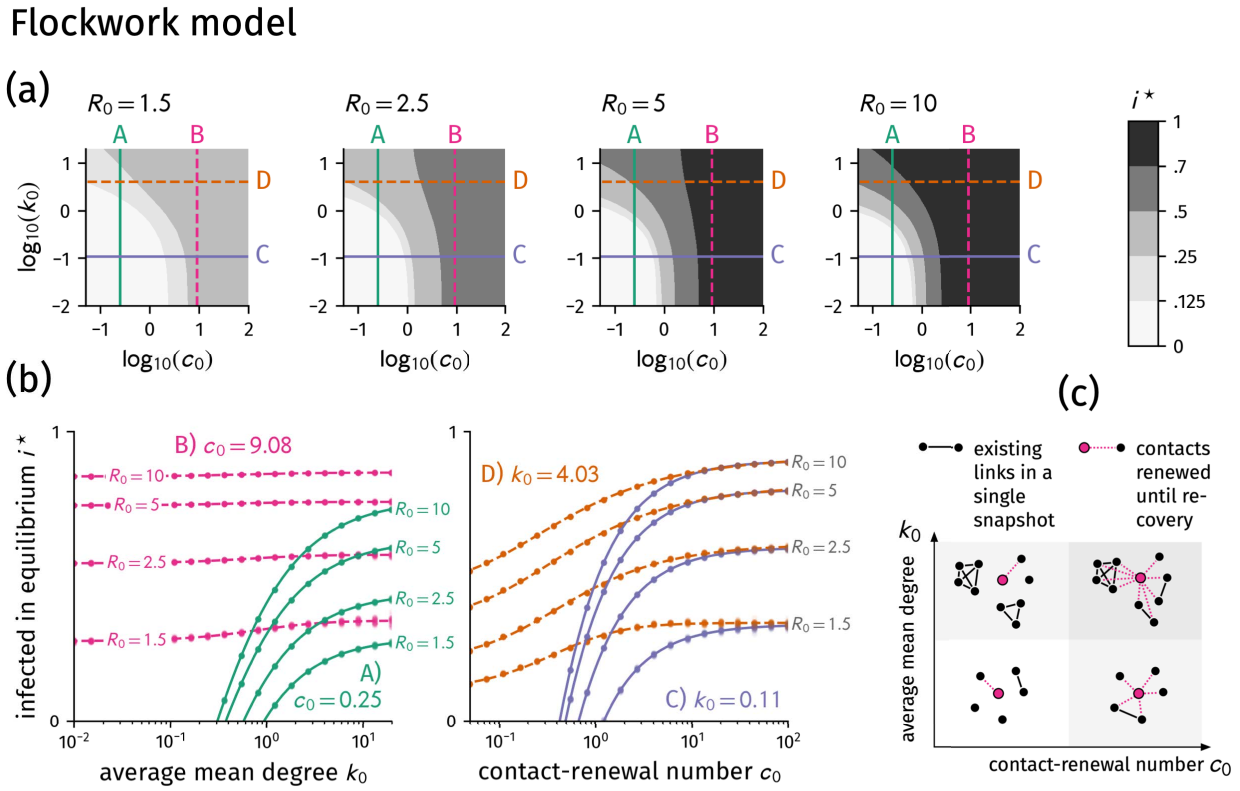


Figure 8.2: Ratio $i^\star$ of infected in equilibrium for Gillespie SIS simulations on the Flockwork model. Here, the average mean degree k_0, contact-renewal number $c_0 = \alpha/\rho$, and the basic reproduction number R_0 are varied. **(a)** For each of the four values of R_0 a disease-free state exists for small contact-renewal numbers and small average mean degree. **(b)** Cuts through the surfaces depicted in (a). For small values of k_0 (c_0, respectively), critical values of c_0 (k_0, respectively) exist, below which the disease dies out. Only for larger values of each respective structural and temporal control parameter, an endemic state is observed. **(c)** Representation of the control parameters. The average mean degree is the mean number of neighbors an average node has at any point in time. The contact-renewal number represents the number of times a node actively reconnects to other nodes before recovering.

an intuitive result: before an infected person can transmit the disease to a new contact, it simply recovers and is very likely isolated, hence no further transmission may happen and the disease eventually dies out.

For small values of the contact-renewal number c_0, large values of the average mean degree k_0 are necessary for a disease to become endemic. In the Flockwork model, a larger mean degree corresponds to the existence of a few groups of larger size while there is still a signficant number of isolated nodes which decreases with increasing k_0. It is hence not surprising that larger groups are necessary for a disease to persist, since isolated infected nodes might recover fast (small c_0 corresponds to a high recovery rate relative to the network dynamics) before they have the opportunity to spread the infection to other nodes.

Both for larger values of k_0 and c_0, respectively, the network structure facilitates a stable endemic state. Focusing on Fig. 8.2b first, one notices that increasing the basic reproduction number shifts the endemic state to a higher number of infected nodes while approximately keeping the shape of the response curve in either k_0 or c_0. For large values of c_0 the system is sufficiently well-mixed such that the mean degree has little influence. Hence, the response curve is flat (see Fig. 8.2b, cuts B). However, for larger values of k_0, the group structure still matters, see the differences between cuts A and B for $k_0 \approx 20$ in Fig. 8.2b, as well as the non-flat form of cuts D. Fur-

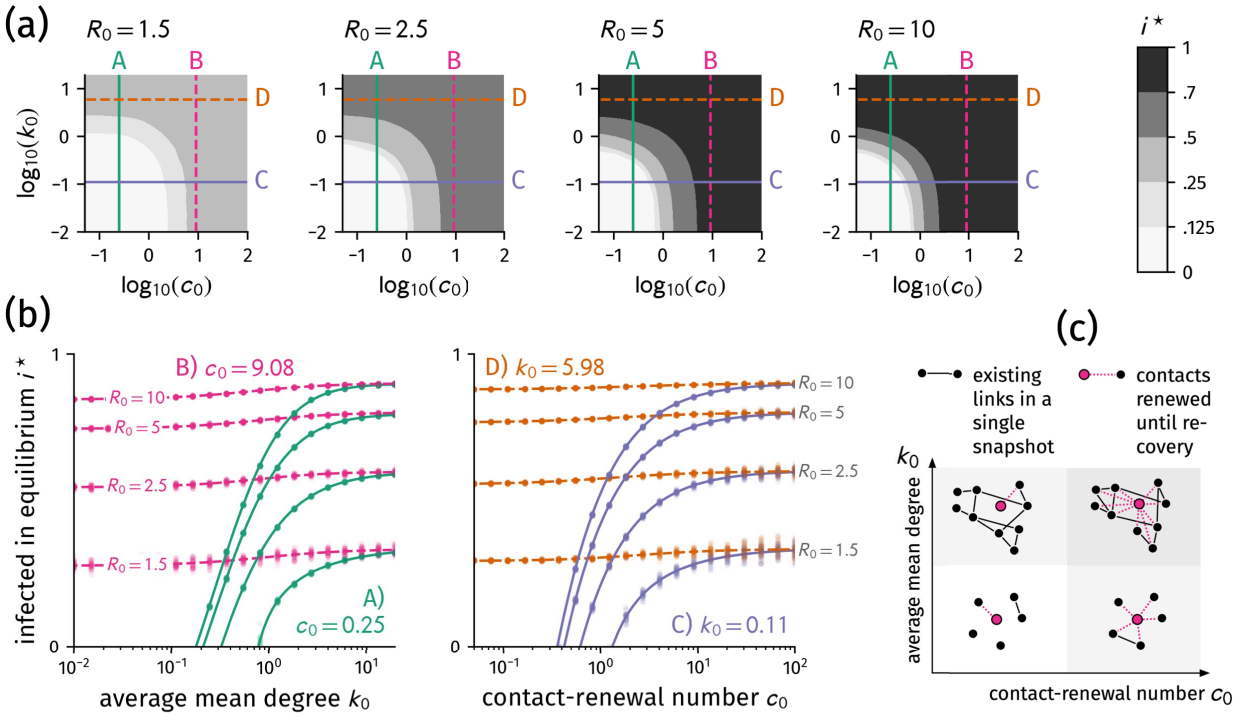


Figure 8.3: Ratio $i^\star$ of infected in equilibrium for Gillespie SIS simulations on the edge activity model. As in Fig. 8.2, the average mean degree k_0, contact-renewal number $c_0 = \alpha/\rho$, and the basic reproduction number R_0 are varied. **(a)** For each of the four values of R_0 a disease-free state exists for small contact-renewal numbers and small average mean degree. The epidemic response surface shows a higher symmetry in between swapping k_0 and c_0 than for the Flockwork model. **(b)** Cuts through the surfaces depicted in (a). For small values of k_0 (c_0, respectively), critical values of c_0 (k_0, respectively) exist, below which the disease dies out. Only for larger values of each respective structural and temporal control parameter, an endemic state is observed. **(c)** Representation of the control parameters. The average mean degree is the mean number of neighbors an average node has at any point in time. The contact-renewal number represents the number of times a node completely changes their neighborhood before recovering.

thermore, a critical mean degree as well as a critical contact-renewal number seems to exist even for high values of the basic reproduction number when low values of the contact-renewal number (the average-mean degree, respectively) are considered (see cuts A and C, respectively). This result will be further discussed in Sec. 8.2.4.

Next, results of the Flockwork model are compared to those of the edge activity model, discussing Fig. 8.3. As for the Flockwork model, one observes in panel (a) that again, for all values of R_0, a stable disease-free state exists for values of $k_0 \lesssim 1$ and $c_0 \lesssim 1$. Since the ER network consists of small disconnected components for $k_0 < 1$, this occurrence can be explained with similar arguments as for the Flockwork model. Contrary to the previous results, however, small values of c_0 and k_0 ($k_0 \approx 1$) can lead to an endemic outbreak. This comes as no surprise because for $k_0 > 1$, an ER network forms a large, connected component of size $\mathcal{O}(N)$ in which the disease can spread while the Flockwork model still consists of small, disconnected components. Finally, for large values of k_0 and c_0, respectively, the network structure facilitates a stable endemic state, similar to the Flockwork model. Focusing on Fig. 8.3b one observes similar behavior as for the Flockwork model with minor differences: In the edge activity model a value of $k_0 \approx 20$ is sufficient to obtain the same endemic state for both small and large values of c_0 (see cuts A and B at $k_0 \approx 20$). In this case, the ER model creates creates well-mixed topologies that give mean-field results (see Fig. 8.3b, cuts B). The same can be observed for high values of c_0 at which

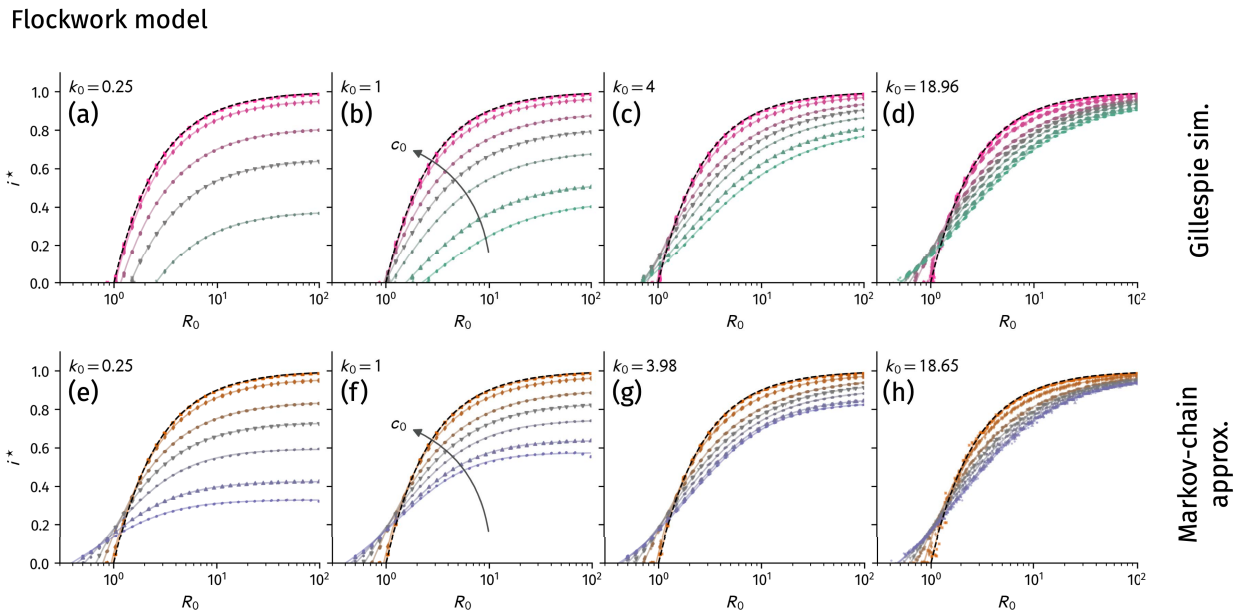


Figure 8.4: Ratio $i^\star$ of infected in equilibrium for SIS simulations on the edge activity model. Here, the mean degree k_0, contact-renewal number $c_0 = \alpha/\rho$, and the basic reproduction number R_0 are varied. Shown are simulation results as well as fits using Eq. (8.2).

the average mean degree does not matter because a node changes its neighborhood sufficiently fast such that the structure essentially equals a well-mixed system (see Fig. 8.3b, cuts D). This is contrary to the result on the Flockwork model where the influence of c_0 is still relevant at $k_0 \approx 20$. Similar to the previous result, one observes a critical parameter $c_{0,c}$ and $k_{0,c}$ even for large values of R_0 if the remaining parameter is sufficiently small (see Fig. 8.3b, cuts A and C).

Next, epidemic response curves are studied in detail as a function of the basic reproduction number R_0. To this end, simulations were performed as described in App. E. In order to fix k_0, the Flockwork reconnection probability P was chosen as $0.1 \leq P \leq 0.95$, then k_0 was calculated using Eq. (7.6). The corresponding parameter value for the edge activity model was consequently given as $p = k_0/(N-1)$. For c_0, the recovery rate ρ per node was fixed as $\rho \in \{0.01, 0.1, 0.5, 1, 2, 5, 10\}$ (in units of $\alpha = 1$) such that $c_0 \in \{0.1, 0.2, 1, 2, 10, 100\}$. The endemic state $i^\star$ was obtained as well as its standard deviation $\mathrm{Std}[i^\star]$ on both the edge activity model as well as the Flockwork model. Furthermore, the individual-based Markov-chain approximation (MCA) equations given by Eqs. (3.26) and (3.27) were integrated on 10 independent continuous-time network simulations of Flockwork and edge activity models for each combination of the parameter values for k_0, c_0, and R_0 with a maximum integration time step[3] of $\Delta t_{\max} = 10^{-2}/\alpha$. Note that for the Markov-chain approximation, networks of $N = 10^3$ were used, which is why some values of k_0 differ insignificantly between the Gillespie and Markov response curves.[4] The integration was stopped either after the demanded simulation time has passed or if a total infection probability

[3] Values obtained for test parameter combinations and $\Delta t_{\max} = 10^{-3}/\alpha$ showed no difference to those obtained with $\Delta t_{\max} = 10^{-2}/\alpha$.

[4] The differences between sample response curves for $N = 10^4$ and $N = 10^3$ were neglible so the computationally cheaper value of $N = 10^3$ was chosen for the larger parameter scan.

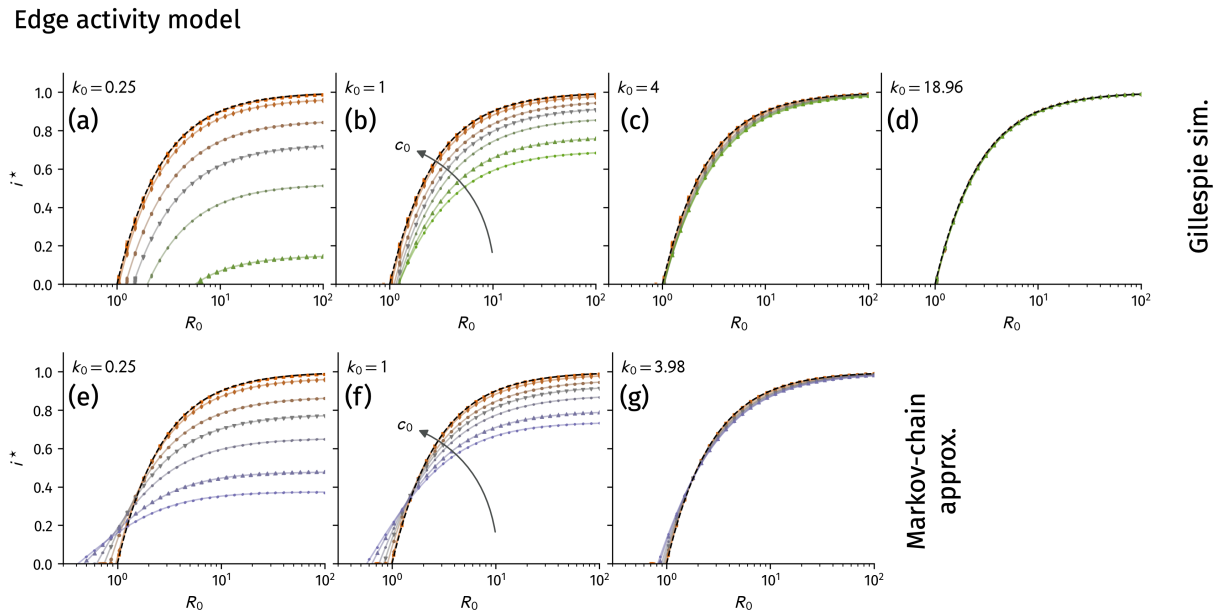


Figure 8.5: Ratio $i^\star$ of infected in equilibrium for SIS simulations on the edge activity model. Here, the mean degree k_0, contact-renewal number $c_0 = \alpha/\rho$, and the basic reproduction number R_0 are varied. Shown are simulation results as well as fits using Eq. (8.2).

of $\sum_{u=1}^{N} \langle i_u \rangle = 10^{-3}$ was reached, in which case the final endemic state was set to $i^\star = 0$.

Results for the Flockwork model are shown in Fig. 8.4a-d. For $k_0 = 0.25$ (panel (a)), the Flockwork model consists mostly of isolated nodes and a few pairs. With high values of c_0, mixing of those pairs is strong enough for the disease to encounter a well-mixed system and hence the response curve approaches the well-mixed case. This behavior is seen for all values of k_0—a high contact-renewal number washes away any single-snapshot structure and thus with $c_0 = 100$ every response curve approaches the mean-field result.

With a decreasing contact-renewal number single-snapshot structure becomes increasingly important. For $k_0 = 0.25$, the asymptotic maximum number of infected people for $R_0 \to \infty$ decreases with decreasing c_0, due to the following reason: At $k_0 = 0.25$, the system consists of many isolated nodes—an increasing recovery rate (decreasing c_0) will lead to more of these isolated nodes recovering before they can be infected again, thus suppressing the endemic state even for large R_0. Furthermore, the epidemic threshold shifts to higher values of R_0, making the system less vulnerable.

Comparing these results to those of the MCA in panel (e), one finds that the maximum number of possible infected people decreases similarly with decreasing contact-renewal number, however, non-zero values of $i^\star$ are reached for small values of c_0, at which no endemic state is observed in the Gillespie simulations in panel (a). Also, the MCA prediction for the epidemic threshold is in direct conflict with those of the simulations. This will be further discussed in Sec. 8.2.3.

With increasing k_0, more nodes can eventually become infected for $R_0 \gg 1$, both, in the simulation results (panel (b)) and the MCA

(panel (f)). When the group size distribution of the Flockwork model becomes increasingly heterogeneous (panels (c) and (d)) and the single-snapshot network structure matters more (low c_0), the degree heterogeneity leads to a decreased epidemic threshold and thus higher susceptibility. For increased group-size heterogeneity, correspondence is increased to the MCA results (panel (g)) and at high values of k_0, the approximation and the simulation strongly correspond (panels (d) and (h)). In this limit, the hetergoeneity in group sizes of a single-snapshot structure seems to reduce the covariance of neighboring node states such that the approximation becomes valid. All panels show that for large R_0 less nodes will be infected as compared to the mean-field result.

The influence of the Flockwork structure is compared to purely random structures by the evaluation of epidemic response curves on the edge activity model (see Fig. 8.5). Panel (a) shows that for a low average mean degree the epidemic response curves qualitatively correspond to the behavior observed in the Flockwork model (Fig.8.4a). This is not surprising as for $k_0 < 1$, the ER model forms small, disconnected components, similar to the Flockwork model. Again, the MCA fails to replicate the Gillespie simulation results for decreasing contact-renewal number c_0 (panel (e)).

With increasing k_0 (panel (b)), the edge activity model possesses a large component of size $\mathcal{O}(N)$ and hence diseases may spread more easily and further than in the corresponding Flockwork model (Fig. 8.4b). While correspondence with the Markov-chain approximation is increased for large values of R_0, the epidemic threshold still shows opposite behavior to that of the simulations (panel (f)).

Further increasing k_0 (panel (c) and (d)), the model quickly resembles a well-mixed system and hence the influence of c_0 vanishes. While for $R_0 \gg 1$ the curves show high correspondence between simulation and approximation (panels (c) and (g)), the epidemic threshold evaluated from the approximation still shows opposite behavior. This will be shown more explicily in the following section. Note that the evaluation of the Markov-chain approximation was omitted for $k_0 \approx 19$ since these showed to be computationally costly. They are expected to show strong correspondence with the mean-field result, much like the simulation results in panel (d). Comparing panel (d) with Fig. 8.4d, the Flockwork model shows a more complex behavior, where the single-snapshot structure of heterogeneous group-size distribution makes the system more susceptible to disease while suppressing the number of people infected in equilibrium. In contrast, the edge activity model simply reflects the well-mixed result.

8.2.3 Epidemic Threshold

Comparing the results displayed in Fig. 8.4(a-d) and Fig. 8.5(a-d), one finds that all curves follow the functional form

$$i^\star = A - \frac{B}{(R_0 - D)^C}. \tag{8.2}$$

This means that for fixed k_0 and c_0, the critical reproduction number $R_{0,c}$ can be computed as

$$R_{0,c} = \left(\frac{B}{A}\right)^{1/C} + D. \tag{8.3}$$

Here, the parameters A, B, C, and D are found using a least-squares fit with initial values $A = 1$, $B = 1$, $C = 1$, and $D = 0$ as these are the corresponding parameters for the mean-field response. For fixed k_0 and c_0, the fits were performed on all Gillespie simulation data pairs of R_0 and $i^\star$ where $i^\star > 0$ and using the simulation values' standard deviation Std$[i^\star]$. The goodness of fit was tested by recognizing the following. Suppose that the data indeed follows Eq. (8.2) with the parameters found by the fitting method. If this is true, one may rescale the data of any simulation using coordinates $x = (R_0 - D)/B^{1/C}$ and $y = (A - i^\star)^{1/C}$ (where A, B, C, and D are the corresponding fit parameters found for fixed k_0 and c_0). Then $y = x^{-1}$ and as such, plotting $\log x$ against $\log y$, all data should collapse to a straight line falling with slope -1. As can be seen in Fig. 8.6, this is indeed the case. Consequently, the epidemic threshold can be found using Eq. (8.3). A second, statistically more rigorous check of Eq. (8.2) is performed in App. E.3 which reveals that the equation tends to overfit the data in certain situations but still reflects it in a fair way. One may use simpler models with only subsets of the parameters A, B, C, and D, however, it was found that doing so does not significantly change the estimation of the epidemic threshold as compared to the full four-parameter model and so the discussion will only focus on fits using Eq. (8.2).

It is found that the epidemic threshold on the Flockwork model displays a rich behavior. As can be seen in Fig. 8.7a, for small values of c_0 and k_0, no epidemic threshold exists, which corresponds to the persistence of disease-free states displayed in Fig. 8.2. Keeping the contact-renewal number c_0 small while increasing the average mean degree k_0, the epidemic threshold decreases, crosses the mean-field result $R_{0,c} = 1$ and decreases towards even smaller values, indicating that the system becomes more susceptible to disease. For small values of c_0, the single-snapshot structure of a temporal network becomes increasingly relevant as an infected node will only 'see' small changes in the network before it recovers. This corresponds to the quenched limit. Large values of k_0 reflect an increasingly heterogeneous group-size distribution which in interplay with low mixing (low c_0), is responsible for the observed decrease in the epidemic threshold.

Increasing the temporal mixing of nodes by increasing the contact-renewal number c_0, the threshold's functional dependence on k_0 remains qualitatively similar, while approaching the mean-field result $R_{0,c} = 1$ for both small and large k_0 until at $c_0 = 100$, the threshold is equal to the well-mixed result: In this case the network is sufficiently annealed such that the local structure does not matter and hence one observes a flat function in Fig. 8.7a.

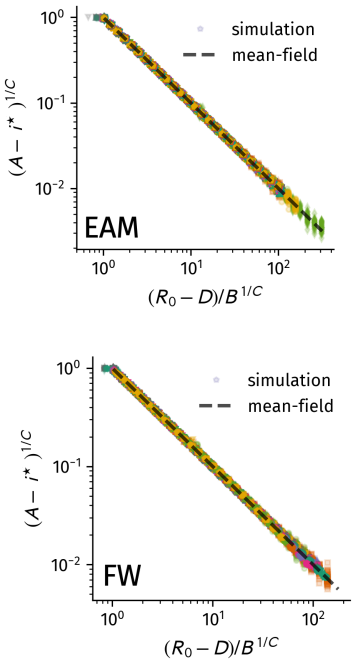


Figure 8.6: Non-zero ratio of infected in equilibrium $i^\star$ as rescaled with the fit parameters of Eq. (8.2). Markers represent all data from the simulations shown in Fig. 8.4. The parameters A, B, C, and D change with varying mean degree k_0 and contact-renewal c_0. For mean-field one has $i^\star = 1 - 1/R_0$ and as such $A = 1$, $B = 1$, $C = 1$, and $D = 0$. **(Top)** simulation and fit results on the edge activity model. **(Bottom)** simulation and fit results on the Flockwork model.

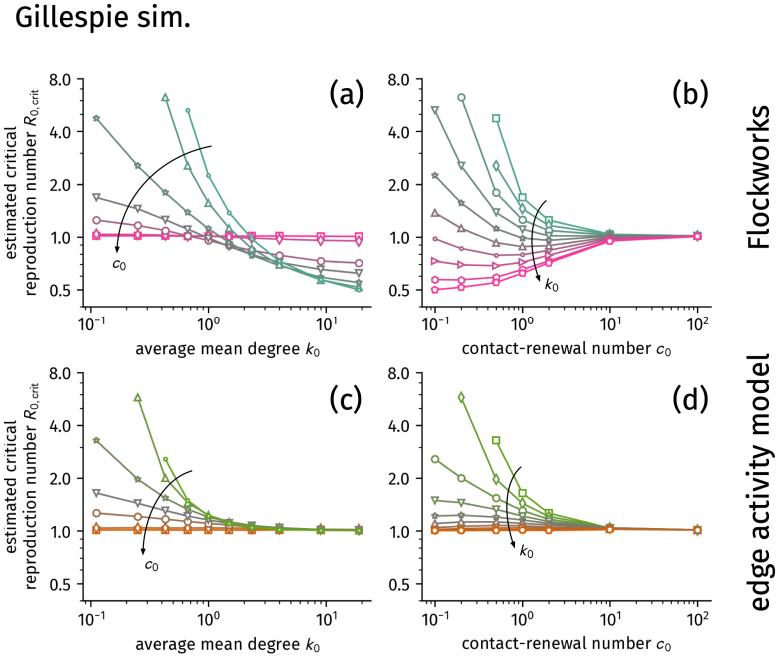


Figure 8.7: Epidemic thresholds estimated from Gillespie simulation results using Eq. (8.3) for both the Flockwork and the edge activity model.

The Flockwork behavior is further displayed in Fig. 8.7b, where the threshold's functional dependence on c_0 is shown while keeping k_0 constant.

In Fig. 8.7c-d, the corresponding behavior of the epidemic threshold on the edge activity model is displayed. As already shown in Fig. 8.3 and similar to the Flockwork model, for small values of k_0 and c_0, the critical reproduction number diverges as no disease can persist in a network of small disconnected components with only low temporal mixing. With both increasing k_0 and c_0, the epidemic threshold decreases towards the mean-field solution $R_{0,c} = 1$. Contrary to the Flockwork model, the epidemic threshold never falls below the mean-field result, and hence, in any case, the system never becomes more susceptible than in its respective well-mixed limits.

Applying fits of Eq. (8.2) to the epidemic response curves obtained by the MCA does not yield statistically proper fits for all parameter combinations. However, comparing the epidemic threshold obtained by the improper fits and Eq. (8.3) with the integrated response curves yields plausible values. Since the aim of this section is to qualitatively compare the results of the MCA and proper continuous-time simulations, the epidemic thresholds estimated by this procedure will suffice.

The results in Fig. 8.8a-b show the estimated epidemic threshold by applying the MCA to continuous-time samples of the Flockwork model. In contrast to the results of Gillespie simulations shown in Fig. 8.7a-b, the epidemic threshold is almost independent on the mean degree k_0. Furthermore, the epidemic threshold exclusively decreases with smaller values of c_0, i.e. when the recovery rate becomes larger. This is in agreement with the result of Speidel *et al.* [45] but in direct conflict with the results from continuous-time simulations.

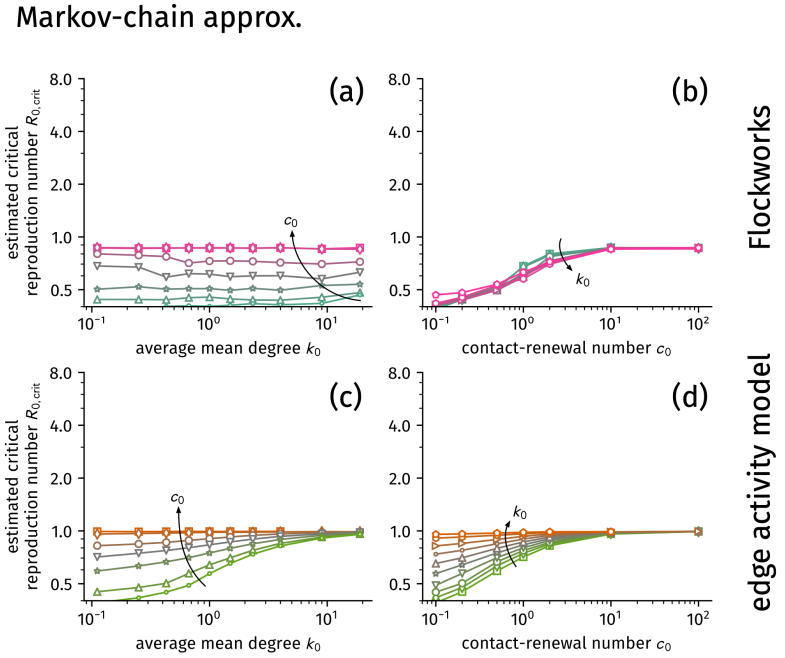


Figure 8.8: Epidemic thresholds estimated from results obtained by the individual-based Markov-chain approximation using Eq. (8.3) for both the Flockwork and the edge activity model.

Only for large values of k_0, the MCA approaches the result of the simulations—supposedly, this is a limit where neighboring-state covariances become negligible as the system is increasingly well-mixed.

For the edge activity model results are shown in Fig. 8.8c-d. For low values of c_0 and k_0, the epidemic threshold is well below the mean-field result, implying that the system is more susceptible to disease. The estimated epidemic threshold then increases when the edge activity model becomes more mixed with increasing c_0 and k_0, approaching the mean-field result from below. Again, this is well within the predictions of Speidel *et al.* but in direct conflict with the results from continuous-time simulations, which showed that the system is never as vulnerable as in the mean-field case.

As was observed in Fig. 8.2a, even for high values of R_0 there exist critical values of c_0 and k_0 below which the disease-free state is reached. In order to illustrate this curiosity more properly, theoretical results for the limiting case of $R_0 \to \infty$ are discussed in the next subsection.

8.2.4 *The Case of Infinite Infectiousness*

Results of the last subsections showed that diseases with high infection rates may not end in an endemic state but rather die out. This happens under the conditions that the network is both very sparse and disconnected, as well as slowly changing compared to the time scale of the disease (as given by a low contact-renewal number c_0, or high recovery rate ρ). To further illustrate this, a limit case for the Flockwork model will be considered in the following where the disease will be infinitely infectious with infection rate $\eta \to \infty$ but is still associated with a finite recovery rate ρ.

An infinite infection rate implies that whenever an infected node connects to a susceptible node, it will infect every susceptible in a group before any other structural event may happen. This way, any reconnection event between an infected and any group of susceptibles is likewise an infection event.[5] In this picture, groups of nodes may represent coarse-grained individuals which can be susceptible and get infected with a rate proportional to the reconnection rate α.

[5] as well as every reconnection event between a susceptible and group of infected, albeit less severe because only the reconnected susceptible node becomes infected.

However, only isolated nodes can recover—any other node that recovers within an infected group will be reinfected instantaneously. Hence, the situation will be expressed in terms of group counts in the following, based on the descriptions of Flockworks in Sec. 7.1.2. The temporal evolution of the number n_g of groups of size g is therefore following Eqs. (7.1.2). At the same time, one may track the number I_g of the groups of size g which are entirely composed of infected nodes. The number of susceptible groups of size g is thus given as $S_g = n_g - I_g$ and the total number of infected nodes in the network is computed as

$$I = \sum_{g=1}^{N} g I_g$$

at all times. Now, events changing the number of infected groups can be categorized as follows.

1. Nodes from infected groups reconnect to other infected nodes (denoted by the superscript (r) in the following). Note that reconnection events happen with active reconnection rate α.

2. Nodes from infected groups detach to become disconnected single nodes (events denoted by the superscript (d) in the following). Disconnection events happen with active disconnection rate β.

3. Nodes from infected groups reconnect to nodes of susceptible groups or vice versa, marking a coarse-grained infection event denoted by the superscript (i).

4. Infected nodes which are isolated may recover, denoted by the superscript (ϱ) in the following.

These events can be summed up to obtain the total temporal evolution equations

$$\partial_t I_g = \partial_t I_g^{(\mathrm{r})} + \partial_t I_g^{(\mathrm{d})} + \partial_t I_g^{(\mathrm{i})} + \partial_t I_g^{(\varrho)},$$

where the single contributions $\partial_t I_g^{(\mathrm{x})}$ are derived in App. C.5. Together with the general group-size evolution Eqs. (7.1) the summed up infection equations

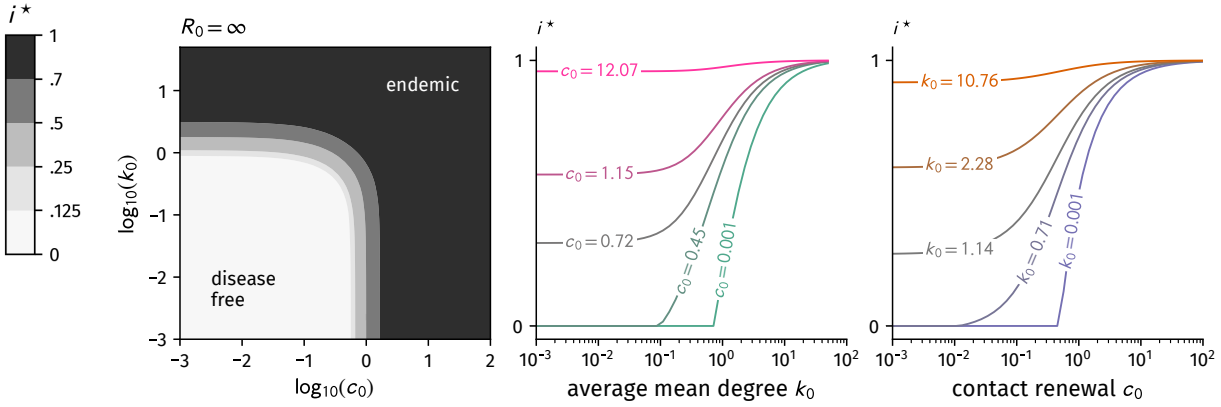


Figure 8.9: Endemic state on the Flockwork model ($N = 100$) for an infinitely infectious disease obtained by integrating Eqs. (8.4) until equilibrium was reached.

$$\partial_t I_1 = -\underbrace{\rho \frac{I_1}{N}}_{\text{recovering}} + \underbrace{\frac{\beta}{N}(I - I_1 + 2I_2)}_{\text{nodes leaving infected groups}} + \underbrace{2\frac{\alpha}{N(N-1)}\Big[I_2(I-2) - I_1(I-1)\Big]}_{\text{reconnections between infected}} + \underbrace{2\frac{\alpha}{N(N-1)}(N-I)\,[I_2 - I_1]}_{\text{reconnections between infected and susceptibles}} \tag{8.4a}$$

$$\begin{aligned}\partial_t I_{g>1} = {} & \frac{\alpha}{N(N-1)}\Big[(g-1)I_{g-1}(I-(g-1)) + \\ & + (g+1)I_{g+1}(I-(g+1)) - 2gI_g\,(I-g)\Big] \\ & + \frac{\beta}{N}\Big[(g+1)I_{g+1} - gI_g\Big] \\ & + \frac{\alpha}{N(N-1)}\Bigg[I(g-1)\Big(n_{g-1} - I_{g-1}\Big) + (N-I) \\ & \times\Big(-2gI_g + (g+1)I_{g+1} + (g-1)I_{g-1}\Big)\Bigg]\end{aligned} \tag{8.4b}$$

approximately describe the average evolution of the whole ensemble of an infinitely infectious SIS process running on a Flockwork model. The validity of these equations is demonstrated in App. E.2 by comparing their temporal evolution to those of simulations with very high infection rate.

Following the simulation setup described therein, the endemic state was evaluated for varying values of k_0 and c_0 by integrating Eqs. (8.4) until an equilibrium was reached. The number of nodes was chosen as $N = 100$ because tests with varying network size showed that the epidemic response curves quickly converge such that the difference between e.g. $N = 50$ and $N = 100$ is already negligible. The results are presented in Fig. 8.9. They reveal that even for an infinitely infectious disease, there exist critical values of the average mean degree and the contact-renewal number below which no disease can become endemic. This result is fairly intuitive: imagine a

highly infectious disease which a person will spread to every person they meet. However, people tend to spend time alone. A disease from which a person recovers before they even meet another person cannot be sustainable.

8.3 Spreading on Real-World Temporal Networks

The results of the last section showed that for a number of Flockwork and edge activity model configurations in dynamic equilibrium, the individual-based Markov-chain approximation (MCA) predicts a decreasing epidemic threshold with increasing recovery rate—while continuous-time simulations show that the epidemic threshold actually increases with increasing recovery rates. Studies based on the MCA found decreasing epidemic thresholds when investigating epidemic spreading on real-world networks. Due to the discrepancy between the results of Gillespie simulations and MCA which arose for models in dynamic equilibrium already, one might wonder about the reliability of the results on real-world networks. Hence, this section aims at investigating continuous-time Gillespie simulations and the MCA on the real-world temporal face-to-face contact data sets DTU, HT09, and HS13, which were introduced in Sec. 2.3.3 and further discussed in Ch. 7.

8.3.1 Methods

The temporal networks will be used as substrates for varyingly paced SIS processes. In Speidel *et al.* [45], the time-scale of a temporal network was chosen to be the duration of a single snapshot. Subsequently, the authors varied the recovery rate in comparison to this length of a snapshot. Since the duration of a single snapshot is a relatively arbitrarily chosen time-scale and varies for different data sets, a more approriate natural time-scale of the data is chosen here. As all data sets discussed here display a circadian rhythm with a period of a single day, their susceptibility to infection will be investigated for the three following time scales: 'slow' diseases ($\rho = 0.125/\mathrm{d}$ and $\rho = 0.25/\mathrm{d}$), 'medium-paced' diseases ($\rho = 0.5/\mathrm{d}$ and $\rho = 1/\mathrm{d}$), and 'fast' diseases ($\rho = 2/\mathrm{d}$ and $\rho = 4/\mathrm{d}$). As discussed in Sec. 2.3.3, the temporal networks are given in a binned manner and therefore discretized in time. They further differ in their observation durations $t_{\max}$ with $t_{\max}^{\mathrm{DTU}} = 7\,\mathrm{d}$, $t_{\max}^{\mathrm{HT09}} \approx 59\,\mathrm{h}$, and $t_{\max}^{\mathrm{HS13}} \approx 111\,\mathrm{h}$. Consequently, their average mean degree k_0 is given by Eq. (8.1) with the corresponding value of $t_{\max}$. Note that the simulation time of the SIS process is usually larger than the respective values of $t_{\max}$ and as such, the temporal network data will be used with periodic boundary conditions in time where $\boldsymbol{A}(t_{\max}) = \boldsymbol{A}(0)$ with $\boldsymbol{A}(t)$ being the adjacency matrix at time t.

The disease's infection rate is given as $\eta = R_0\rho/k_0$ for varying R_0. For each parameter combination, $N_{\mathrm{meas}} = 100$ independent simulations were perfomed. After an initial equilibration phase of duration

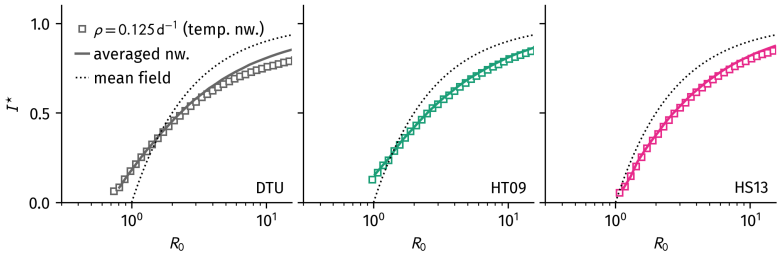


Figure 8.10: Epidemic response curves from Gillespie simulations on the averaged temporal networks and the temporally resolved networks for a low recovery rate (slow disease) for DTU, HT09, and HS13 data sets. The simulation results obtained on the averaged network correspond to the simulation results on the temporal network in most cases (small differences for high infection rates in the DTU data set).

t_{eq}, the number of infected nodes was sampled every $\Delta t_{\mathrm{sample}} = 1\,\mathrm{h}$ over a duration of $t_{\mathrm{meas}} = M\Delta t_{\mathrm{sample}}$.[6] Those n simulations which did not reach the absorbing state $i^\star = 0$ were further analyzed. If at least $n \geq 10$ of those simulations reached a state $i^\star > 0$, the average time-dependent curve was evaluated as an average over all non-zero simulations where $i^{(\sigma)}(t)$ denotes the σ-th simulation to obtain

$$i^\star(t) = \frac{1}{n}\sum_{s=1}^{n} i^{(\sigma)}(t). \tag{8.5}$$

The mean average endemic state was then computed as

$$i^\star = \frac{1}{M+1}\sum_{r=0}^{M} i^\star(t_{\mathrm{eq}} + r\Delta t). \tag{8.6}$$

More details concerning Gillespie simulations and the MCA evaluations on temporal networks can be found in App. E.4.

[6] If not indicated otherwise, M was chosen such that $t_{\mathrm{meas}} = 70\,\mathrm{d}$.

8.3.2 *Averaged Networks*

In theory, the longer a node is infected, the less the single-snapshot structure of the network should matter. Instead, the node should see an averaged network. In order to test this hypothesis, the result of the lowest recovery rate Gillespie simulation on temporal networks will be compared to the result of a Gillespie simulation on the averaged weighted network.

As can be seen in Fig. 8.10, simulating an SIS process on all data sets with low recovery rate of $\rho = (1/8)\mathrm{d}$ yields an epidemic response curve corresponding to the response curve of their averaged network. The difference between the results for temporal network and the average network slightly increases at high values of R_0. This comes from the fact that an average infected node is infected for $t = 1/\rho = 8\,\mathrm{d}$ which means that it effectively only experiences a single period of $t_{\mathrm{max}} = 7\mathrm{d}$, while in the HT09 and HS13 data sets, such a node will see approximately 1.5 to 2 full periods. This analysis confirms again that the annealed limit $\rho \to 0$ corresponds to epidemic spreading on the corresponding static network.

8.3.3 *The Influence of Circadian Activity*

Real-world temporal networks show strong circadian activity. The influence this activity has on epidemic spreading is not well-studied,

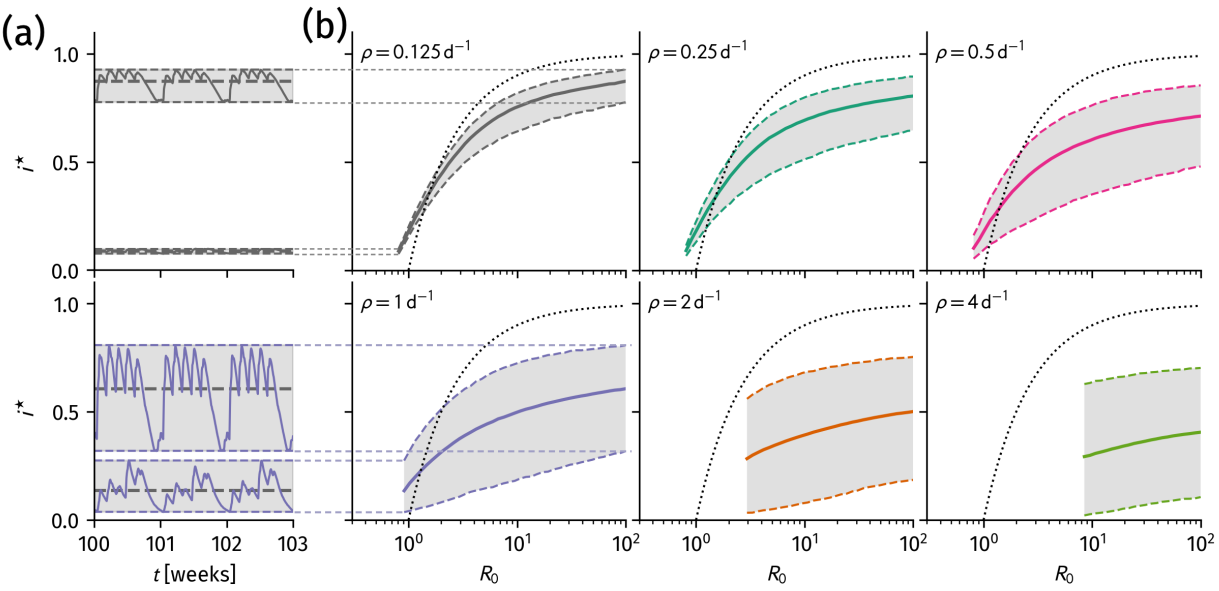


Figure 8.11: The influence of circadian activity on epidemic response curves and epidemic thresholds. Since the networks investigated here become sparse and vary slowly during the night (or during the weekend), the endemic state potentially decreases strongly during those times. **(a)** Temporally varying endemic states averaged over $n \geq 10$ simulations in the observation period and the extent of their variation marked as grey bands. **(b)** Epidemic response curves. Solid lines represent the average endemic state Eq. (8.6) and grey bands mark the extent of its variation. When the minimal number of infected $i^\star_{\mathrm{min}}$ (lower dashed lines) approaches $i^\star_{\mathrm{min}} = 0$, the disease becomes unsustainable and dies out. The average endemic state is potentially much larger than the minimal endemic state and therefore, sudden jumps in the order parameter $i^\star$ occur.

and therefore some results are presented here. Since all data sets show similar circadian activity, it will suffice to analyze the DTU set as an example. Nevertheless, results for all data sets and Flockwork surrogate networks are presented in App. E.4.2. The results of the Gillespie simulations are shown in Fig. 8.11.

For each value of R_0 and ρ, the average number of infected $i^\star(t)$ was measured as a function of time over $t_{\mathrm{meas}} = 3 \times 7\,\mathrm{d}$ after the equilibration period of $t_{\mathrm{eq}} = 100 \times 7\,\mathrm{d}$. Over this period, the endemic state $i^\star$ was found using Eq. (8.6). Furthermore, $i^\star_{\mathrm{min}}$ and $i^\star_{\mathrm{max}}$ were found as the minimum and maximum values, respectively, of $i^\star(t)$ over this period of time. In Fig. 8.11a-b, solid lines represent the average endemic state Eq. (8.6) while the grey band enclosed by dashed lines show the extent of the endemic state's variation over the measurement period. As one may see, for all investigated recovery rates, the curves show broad bands, indicating strong variation of the endemic state over time. For slow diseases, this variation decreases close to the epidemic threshold such that it suffices to observe the average endemic state to deduce the epidemic threshold. However, already for medium-paced diseases, the influence of circadian activity becomes stronger, leading to sudden jumps from $i^\star = 0$ to $i^\star > 0$ with increasing R_0. Usually, jumps like these are attributed to fluctuations and become less prominent with larger system sizes. Another method to avoid these jumps is to simulate the reactive process using the quasi-stationary state method as discussed in Sec. 3.2.3. Yet, the networks discussed here are not in a stationary state. Rather, as discussed in Sec. 7.2.2, face-to-face contact systems can be intrepreted to follow a trajectory in the (c_0, k_0) state space. Investigating a single disease of infection rate $\eta_0 = R_0\rho/k_0$, this trajectory will lie on a plane defined by $\eta = \eta_0$ in the three-dimensional epidemic (c_0, k_0, η) state space (see Fig. 8.12). As established for model networks in the

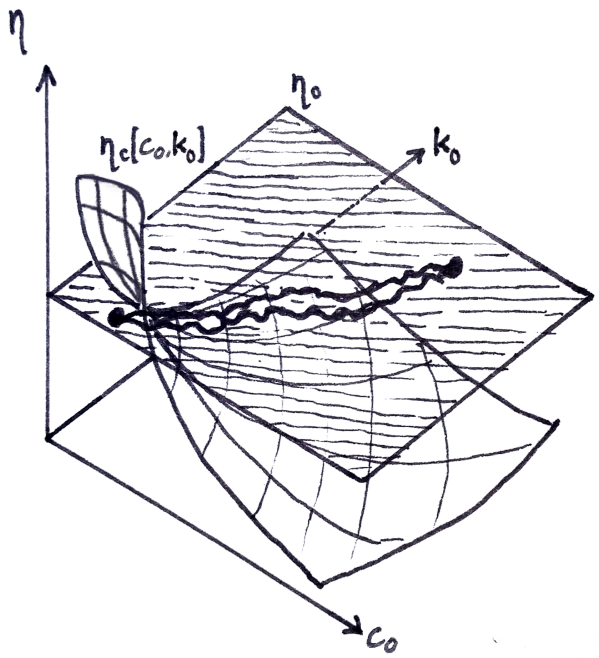


Figure 8.12: As discussed in Sec. 7.2.2, temporal networks can be interpreted to follow a trajectory in the (c_0, k_0) state space. Considering the severity of a disease as defined by its infection rate, this trajectory lies on a plane η_0 in the (c_0, k_0, η) state space. In this epidemic state space, the critical surface $\eta_c(c_0, k_0)$ separates the regions where the system is disease-free and where it is endemic, as was found in the last section. The trajectory of a temporal network may cross this critical surface to enter the absorbing region. The system is more likely to actually reach the disease-free state the longer it stays in this region (with longer referring to time in units of the node recovery rate ρ^{-1}).

last section, a critical surface defined by $i^\star(c_0, k_0, \eta_c) = 0$ separates the disease-free and endemic states. When the trajectory enters the disease-free region of this state space, the disease can die out if the system remains there for long enough. The following considerations might explain why such jumps were not discussed in other studies yet:

For time-dependent quasi-stationary state simulations, as they have been used for instance by Speidel *et al.*, configurations ending in the absorbing state are replaced with non-zero infected configurations at similar structures (similar times in the observation period). Hence, simulations obtaining the 'quasi-stationary' endemic state $i^\star(t)$ will remain active even on the disease-free side of the critical surface until the trajectory enters more infectious regions again where $i^\star(t) > 0$. When using the average endemic state $i^\star$ as an order parameter, quasi-stationary algorithms will therefore predict a non-zero endemic state, when in reality, the disease might have well died out in the slow/sparse region. Thus, quasi-stationary algorithms might underestimate the epidemic threshold, at least in finite systems, when $i^\star$ is chosen as the order parameter. In contrast, one could choose the minimal endemic state $i^\star_{\text{min}}$ (minimal during the temporal network observation period) as the order parameter. In Fig. 8.13, a comparison to the corresponding observables obtained by the Markov-chain approximation (MCA) method on the DTU data is shown. For increasing recovery rates, the MCA prediction of the minimal number of infected nodes $i^\star_{\text{min}}$ increasingly differs from the simulation result, approaching larger values than obtained by simulations. In simulations, vanishing values of the minimal endemic

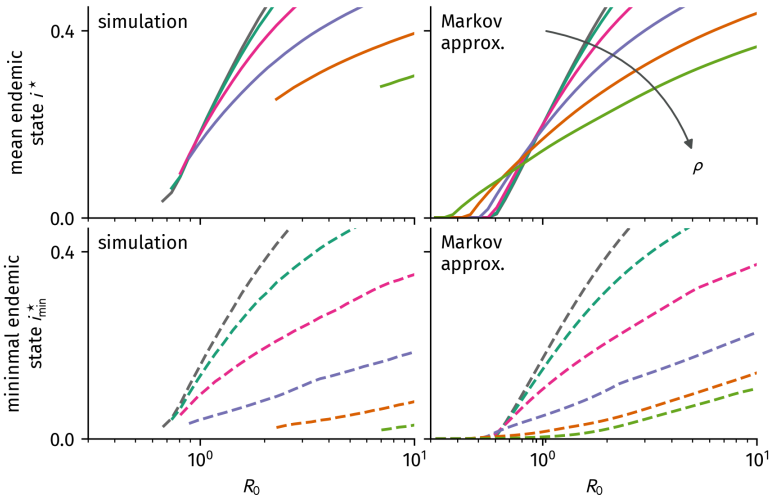


Figure 8.13: Mean endemic state $i^\star$ and minimal endemic state $i^\star_{\min} = \min\{i^\star(t)|t_{\text{eq}} \leq t \leq t_{\text{eq}} + t_{\text{sim}}\}$ for varying recovery rates on the DTU data, comparing Gillespie simulations and results from the Markov-chain approximation (MCA). (Top row) solid lines represent the average endemic state Eq. (8.6) and (bottom row) dashed lines represent the minimal number of infected $i^\star_{\min}$ during the observation period. (Left column) response curves obtained by Gillespie simulations. (Right column) response curves obtained by the MCA. When the minimal number of infected $i^\star_{\min}$ (lower dashed lines) approaches $i^\star_{\min} = 0$, the disease becomes unsustainable and dies out. While traditionally, sudden jumps in the epidemic response curves are associated with fluctuations in the average endemic state, they should be associated with fluctuations in the minimal number of infected $i^\star_{\min}$. A side-by-side comparison of both is shown in the Appendix, Fig. E.4.

state lead to large sudden jumps in the order parameter $i^\star$, which cannot be reproduced by the MCA, which wrongly overestimates the minimal endemic state. For HT09 and HS13 strong differences arise even for lower recovery rates (shown in App. E.4.2). It is therefore questionable whether the MCA predicts correct epidemic thresholds, at least in systems of finite size.

To sum up, when investigating circadially varying structures, focus should be given to the minimal number of infected during a measuring period $i^\star_{\min}$. As soon as the configuration reaches a state where $i^\star_{\min} = 0$ it will inevitably die out. Hence, sudden jumps to large values of the average endemic state $i^\star$ are suspected to be characteristic at the epidemic threshold when considering large recovery rates on circadially varying temporal networks. Since so far, no continuous-time formulation of a quasi-stationary simulation algorithm exists, the exact value of the epidemic threshold cannot be deduced from the simulations here.

8.3.4 *Comparison of Gillespie Simulations and Markov-Chain Approximation*

Here it will be investigated whether the MCA is a valid approach to study SIS spreading on the temporal networks and recovery rates introduced above by comparing epidemic response curves. The results are shown in Fig. 8.14.

The MCA results in similar response curves for low recovery rates (slow diseases) for the DTU data set but produces increasingly different results for higher recovery rates. Consistently, the epidemic threshold obtained by the approximation method slightly decreases with increasing recovery rate (comparable to the claim by Speidel *et al.* [45], see Fig. 8.13, upper right panel). This is, however, in contrast to some of the results of the Gillespie simulations, especially for the HS13 dataset where, most likely, the opposite is true.

This comes as no surprise because the MCA is only valid in sys-

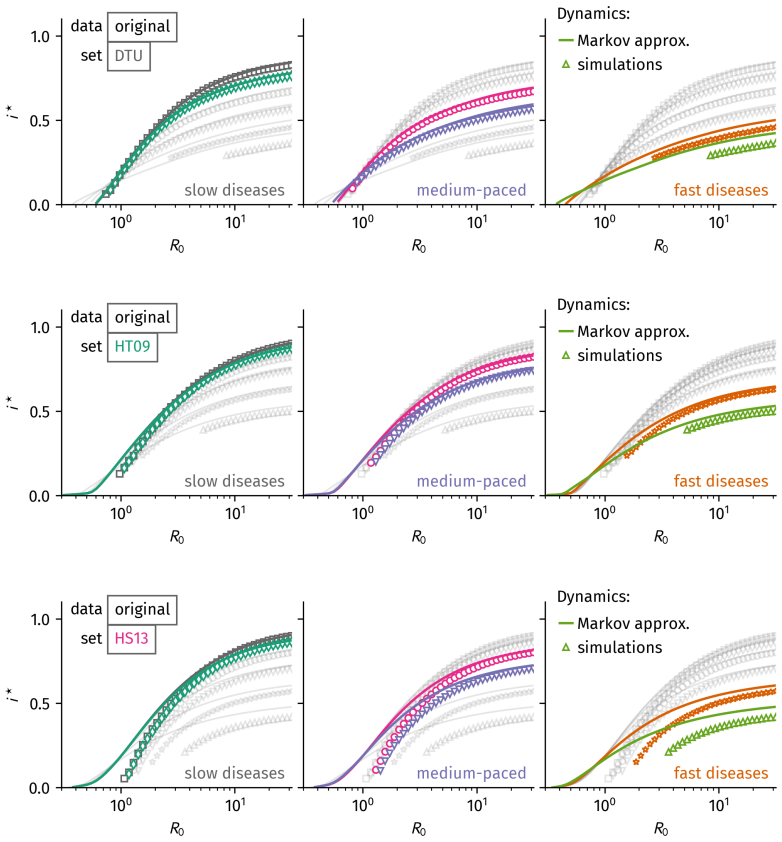


Figure 8.14: Epidemic response curves for DTU, HT09, and HS13, comparing results from Gillespie simulations to those of the Markov-chain approximation (MCA). While there is strong correspondence for slow diseases on the DTU data set, there is less correspondence for all other cases, especially in the case of low reproduction numbers R_0 and fast diseases. The MCA predicts decreasing epidemic thresholds with increasing recovery rates while the continuous-time simulations show opposite behavior.

tems where the infectious states of neighbors are uncorrelated. Such an approximation may be reasonable for low recovery rates where nodes frequently change their neighbors before recovering and denser networks, both representing cases closer to well-mixed systems. The assumption is, however, not valid for shorter infection times and structures which resemble lattices (as is the case for the high school network HS13 where people tend to spend time preferrably with their classmates). The approximation is also invalid for sparse networks—however, all of the contact networks are sparse during the night (so roughly half of the time). An illustration for how the MCA drastically underestimates infection state covariance in a modified SIS process on a single pair of nodes is shown in App. E.5.

As discussed in the last subsection, the epidemic response curves show sudden jumps in the order parameter $i^\star$, especially for fast diseases, due to the finite system size and influence of circadian activity. Therefore, definite statements about the increase/decrease of the epidemic threshold cannot be made. In order to make definitive statements about increasing or decreasing thresholds, a quasi-stationary state algorithm monitoring the minimal number of infected nodes per period should be used in future studies, as discussed in the last subsection.

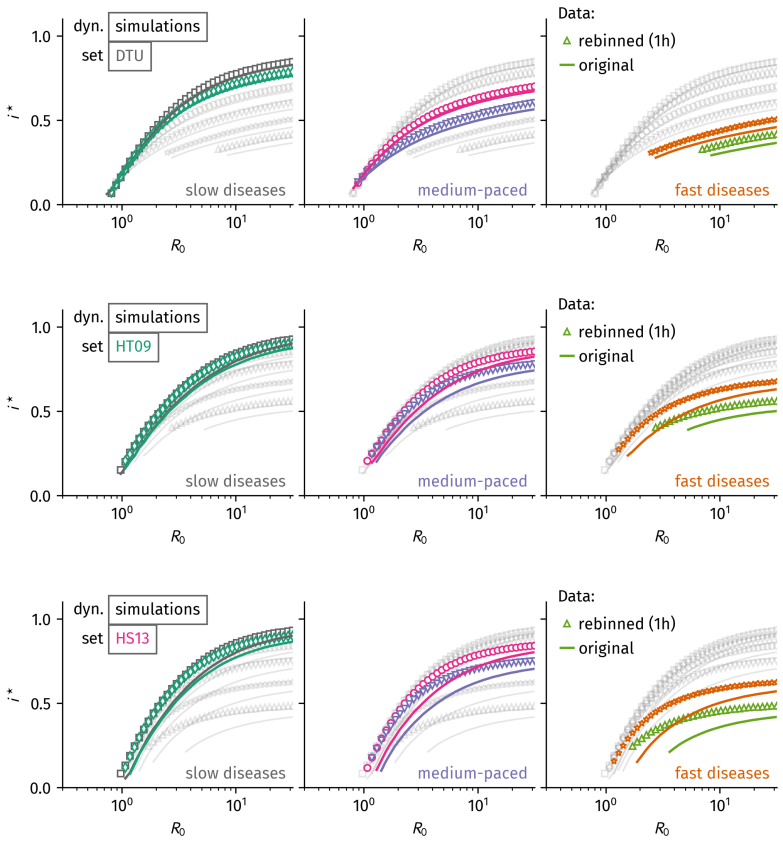


Figure 8.15: Epidemic response curves for DTU, HT09, and HS13, comparing results from Gillespie simulations on both original and rebinned data. Rebinning leads to an increased endemic state and decreased epidemic thresholds.

8.3.5 *The Influence of Rebinning Temporal Network Data*

In the previously discussed studies [45, 56, 91], temporal network data was rebinned to larger aggregation windows and sometimes further manipulated before reactive process simulations were compared to the results of the MCA. As was shown and discussed in Ch. 7 this transformation drastically changes the structure of temporal networks regarding their single snapshots (for instance the group-size distribution, cf. Fig. 7.16). Hence, it will be investigated here how the epidemic response curves change with increased aggregation duration as compared to simulations on the original networks. To this end, every network is rebinned to $\Delta t = 1\,\mathrm{h}$ and the average mean degree is recalculated to properly gauge R_0. Subsequently, Gillespie simulations are run on the rebinned networks and compared to the corresponding response curves on the original data. The results are shown in Fig. 8.15.

Rebinning the networks shifts the response curves to higher values, albeit not drastically for slow and medium-paced diseases on HT09 and DTU. They are, however, increasingly different for HS13 and in general for higher recovery rates. One may still observe the sudden transition to the endemic state for high recovery rates—the curves jump to a similar value of $i^\star$ for the same recovery rate, however at different values of R_0. This reinforces the notion that the jumping behavior is induced by the circadian activity pattern dis-

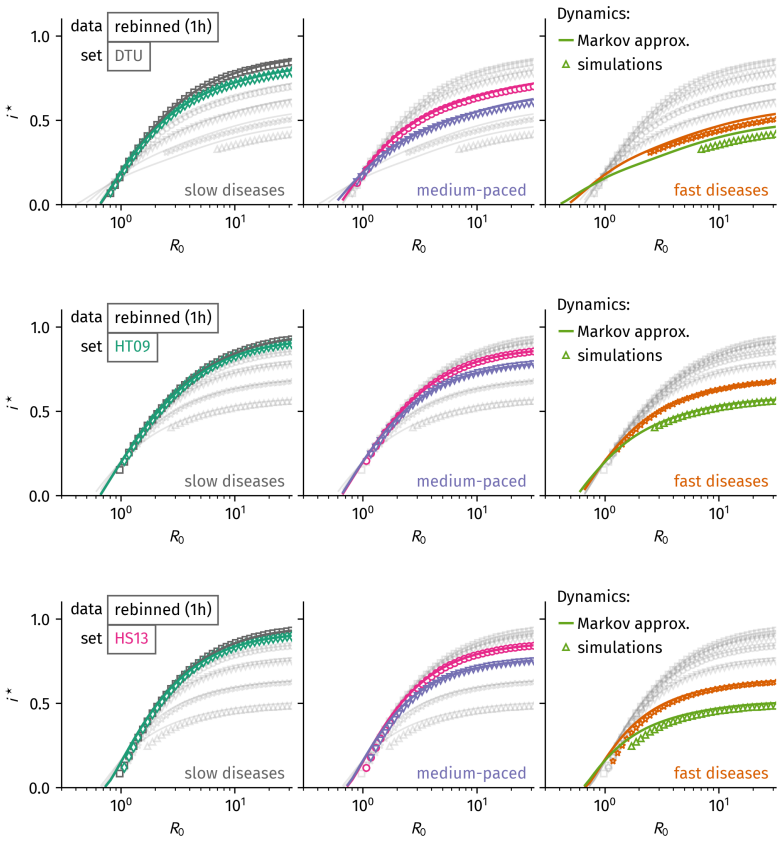


Figure 8.16: Epidemic response curves for rebinned data of DTU, HT09, and HS13, comparing results from Gillespie simulations to those of the Markov-chain approximation. The correspondence between the curves of both methods is strongly improved compared to simulations on the original data.

played in all the data sets, which is not influenced too strongly by rebinning with $\Delta t = 1\,\mathrm{h}$. In most of the cases the epidemic threshold appears to decrease after rebinning.

One may conclude that it is in general not a safe assumption that rebinning will not change the epidemic response strongly even if the new aggregation duration is much lower than the average recovery time.

In [91], Valdano *et al.* investigate the influence of rebinning on epidemic thresholds. They do so, however, based on the MCA only, neglecting results from actual simulations. In the following it will be tested whether the MCA yields better results on rebinned temporal networks in comparison to accurate continuous-time simulations. This comparison is shown in Fig. 8.16.

Even though the MCA still appears to underestimate the epidemic threshold, there is much stronger correspondence between the curves of both methods on rebinned data. A possible reason is that aggregation decreases sparsity and therefore state covariance between neighbors, making the MCA a more valid approach. However, the investigated substrate for a disease to spread is essentially different from the original structure.

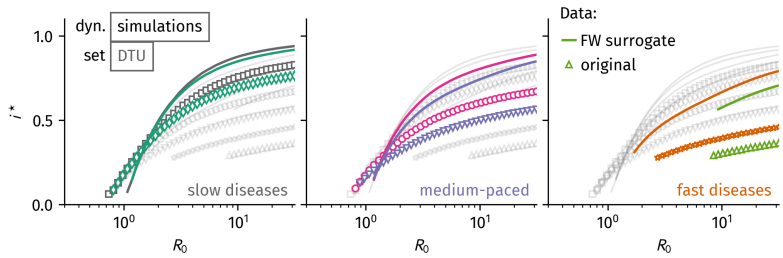


Figure 8.17: Epidemic response curves for simulations on Flockwork surrogate data generated to resemble the single-snapshot structure and circadian activity variation of the DTU data set, compared to results on original data. Response curves on Flockwork surrogate and original data strongly differ, most likely influenced by strong random mixing behavior of nodes in the Flockwork model. Both systems display sudden jumps in the response curves for fast diseases.

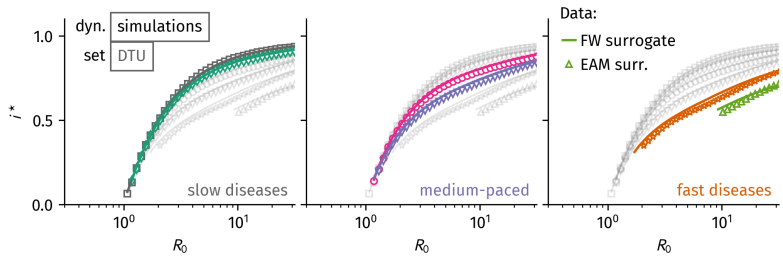


Figure 8.18: Epidemic response curves for surrogate data generated from the Flockwork and edge activity models to mimick the DTU data set and compared to the result of the corresponding Markov-chain approximation. The response curves show strong correspondence, indicating that their structural differences of single-snapshots in dense networks is negligible for the recovery rates investigated here.

8.3.6 Surrogate Networks

After modeling real-world temporal networks based on the Flockwork model in the last chapter, one may raise the question whether the shape of the epidemic response curves can be explained solely by considering group-structured topologies and circadian activity. To this end, a circadially varying Flockwork surrogate network resembling the DTU data set is investigated as the substrate for an SIS model. In the Flockwork model, the averaged network has no structure but resembles a completely connected network, because when a node reconnects, each node has equal probability to be chosen as a neighbor. The resulting response curves based on Gillespie simulations are shown in Fig. 8.17. Results for the other data sets are shown in App. E.4.2.

The Flockwork simulations do not represent the original response curves in a satisfying manner. This is not surprising for the lower recovery rates: a single node 'sees' a rather averaged network which, in the Flockwork case, is very different from the richer structures of the averaged real-world network. The only replicated effect is the sudden jump of the response curve at higher recovery rates, which is most likely induced by circadian activity, as discussed above.

This assumption is tested by comparing the results of the Flockwork surrogate to the results of a corresponding edge activity model surrogate generated from the same circadian activity rates as the Flockwork surrogate with results shown in Fig. 8.18. Even though both models have a drastically different single-snapshot structure at times where k_0 is large, this difference does not matter for low recovery rates. The shape of the response curves is solely dominated by the similar circadian activity pattern and the corresponding averaged network, which resembles a completely connected graph for both

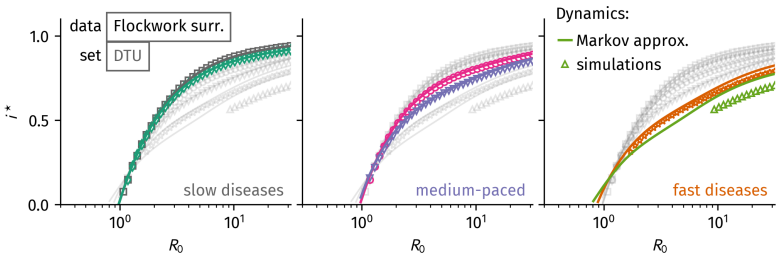


Figure 8.19: Epidemic response curves for surrogate data generated from the Flockwork model for the DTU data set, both from simulations and the individual-based Markov-chain approximation. There is high correspondence between the curves, however, the jumping behavior at the epidemic threshold is not reproduced.

models. An increasing difference can be seen for higher recovery rates where a single infected node sees more of the single-snapshot structure before it recovers. Nevertheless, the sudden jumps in the response curves are positioned at similar values of R_0, again implying that circadially varying activity is the cause.

Since Flockwork networks resemble completely connected networks for low recovery rates and the covariance of neighboring infection states is low in those networks, it will be tested whether the MCA is valid for the Flockwork surrogate networks. In Fig. 8.19, the results of the approximation are compared to those of Gillespie simulations.

Indeed, the Markov integration method satisfyingly corresponds to the response curves of the Gillespie simulations, besides an increasing difference for high recovery rates and the sudden jumps in the order parameter, which are not reproduced. Since local structures play more prominent roles during short recovery times (high recovery rates), the networks' sparsity at night supposedly induces high covariances between infectious states which leads to discrepancies between the individual-based approximation and the Gillespie simulations and an underestimation of the epidemic threshold in the approximation. After comparing simulation and approximation results for all Flockwork surrogate data (see Fig. E.6 and Fig. E.7 in App. E.4.2), the MCA seems to predict a slightly decreasing epidemic threshold, which might, again, not be confirmed by the simulation. However, rather large jumps in the epidemic response curves obfuscate a clear conclusion.

8.4 Summary and Discussion

In this chapter, the SIS model was studied on temporal contact networks, where the influence of single-snapshot temporal structure was increased by varying the recovery rate of a single node from $\rho \to 0$ to $\rho \gg 0$, i.e. from a situation where a single infected node experiences the averaged temporal network before recovering (annealed limit) to a situation where few changes of the network structure were relevant (quenched limit). Former studies stated that the epidemic threshold decreases with increasing recovery rate, making systems more susceptible to diseases when its temporal variation is of increased importance. However, these statements were based on

studying reactive processes and the individual-based Markov-chain approximation (MCA). Since the reactive process can be different from the basic Poissonian SIS process, this chapter mainly studied results based on exact continuous-time simulations and compared these results to those of the MCA.

In a first investigation, the influence of a temporal network's single-snapshot structure on epidemic spreading was discussed. To this end, two basic random temporal network models were chosen as substrates for a disease to spread: the Flockwork model, where the network consists of disjoint, densely connect groups and isolated nodes and the edge activity model, where each single snapshot is a sample from the Erdős–Rényi (ER) network model. Temporal variation is introduced via an active reconnection rate α which counts the number of times a node connects to a new neighborhood per unit time. Both models resemble a completely connected graph in the annealed limit.

In order to fix the time scale, the *contact-renewal number* $c_0 = \alpha/\rho$ was introduced, which quanitifies the number of times a node changes its neighborhood before it recovers. While the average mean degree k_0 quantifies the *structural* mixing of a single node, this new control parameter therefore quantifies the *temporal* mixing of a node.

For the simpler edge activity model, it was found that large values of either c_0 and k_0 let the system behave like a mean-field system: nodes 'see' a well-mixed and average medium. Consequently, the MCA correctly predicts the epidemic response curve, allegedly because the covariance between neighboring infection states becomes negligible. However, decreasing c_0 and k_0, the epidemic threshold increases, making the system less susceptible to diseases. This is in direct contrast to the result of the MCA which predicts a decrease in the epidemic threshold.

For the Flockwork model, it was shown analytically that for $R_0 \to \infty$, there exist critical values of k_0 and c_0, respectively, below which no disease can be sustainable. This is an intuitive result: A disease which is unusally infective, but nevertheless infected people recover incredibly fast, does not exist. Concerning finite basic reproduction numbers $R_0 < \infty$, the MCA predicts the epidemic response well for large values of k_0, when the group-size distribution becomes heterogeneous. In this case, both MCA and Gillespie simulations show a decreasing epidemic threshold with increasing recovery rate. However, the MCA directly contradicts the simulation results for lower values of k_0 when groups become small and the group-size distribution homogeneous: Simulations show that for those structures, an increased recovery rate yields increased epidemic thresholds while the MCA still predicts decreasing epidemic thresholds.

These results implicate that new theoretical advances are necessary to accurately predict the epidemic response in sparse model networks in continuous time. Current theoretical approaches are based on the MCA [45, 56, 57, 91]. In [57], Valdano *et al.* argue that this approach can yield exact results on the epidemic threshold

even for continuous-time temporal networks. This is, however, not the case: the exact continuous-time simulations based on continuous-time temporal network models shown in this chapter revealed that the results obtained by the MCA yield contradictory behavior. While in the studies mentioned above, the theoretical approaches were indeed confirmed by simulation results, these simulations were performed with the reactive process: time is discretized and during each time-step of length Δt, each infected infects each of its susceptible neighbors with probability $\eta\Delta t$ and afterwards recovers with probability $\rho\Delta t$. A node can recover and be re-infected within a single time step. This process is only equal to the exact continuous-time process if the probabilities $\eta\Delta t$ and $\rho\Delta t$ are sufficiently small such that during a single time step of Δt either no or maximally one single event happens. Only then is the condition for a Poisson process met (see App. D). This means that if one takes the continuous-time limit $\Delta t \to 0$, one should simulate the reactive processes with probabilities $\eta\Delta t \to 0$ and $\rho\Delta t \to 0$. However, it seems to be practice to simulate the reactive process with finite probabilities even approaching unity. In this case, multiple events may happen at once and hence the simulation results will not follow those of an exact continuous-time Poisson process. A similar critique of simulation and theoretical approaches on *static* networks was already given in [52].

Furthermore, in the studies cited above, theory and simulations were tested on randomly generated discrete-time temporal network models: Adjacency matrices last for a single time-step of duration Δt, after which a new adjacency matrix is generated, completely independent from the last adjacency matrix (see Sec. 2.4.3). The Flockwork model and the edge activity model are therefore valuable additions to the set of temporal network models. Since they are based on Poisson processes, they are naturally formulated in continuous time and hence prove to be ideal candidates for simulations of continuous-time dynamic processes.

While results for the Poissonian Flockwork model were obtained here, real-world networks are not of constant activity rates but might be interpreted to follow structural trajectories of equilibrium states in the (c_0, k_0)-plane over time (as shown in the last chapter). The Flockwork model was designed to assume the single-snapshot structure of real-world temporal networks. At times where there is a lot of activity and larger groups, the MCA can be expected to correctly predict the epidemic response as discussed above, but it might fail to do so for low values of k_0. Therefore, exact continuous-time simulations on real-world network data were performed and compared to the corresponding results obtained by the MCA.

It was found that the MCA accurately predicts epidemic response curves for low recovery rates on the DTU data set but fails to describe the simulation results for the HT09 and HS13 data sets. Even for low recovery rates, differences arise close to the epidemic threshold. Continuous-time simulations with low recovery rate yielded similar results on both the average as well as the temporally resolved

network. Therefore, the MCA is already an invalid approach in the annealed, static limit, at least for the HT09 and HS13 data sets.

Furthermore, the epidemic threshold appeared to decrease with increasing recovery rate, contradicting the results of the MCA. However, due to the limited size of the networks, the exact position of the epidemic thresholds could not be determined. Investigating the epidemic thresholds it was found that circadian variation in sparsity and activity may lead to sudden jumps in the epidemic response curves in those finite systems: while the number of infected nodes averaged over the observation period might still be large, the average *minimal* number of infected during a certain observation period is approaching the absorbing state. These differences are induced by nodes being mostly isolated during the night—if the recovery rate is large enough, all nodes may recover during the night without being able to pass the infection further. In the thermodynamic limit $N \to \infty$, this behavior should disappear: While the number of infected nodes decays exponentially during the night, it only reaches the disease-free state asymptotically for $t \to \infty$. After night has passed, however, newly built contacts may yield a surge in infected people. Consequently, the MCA, which assumes the thermodynamic limit, does not yield epidemic response curves with sudden jumps. Nevertheless, the MCA does not correctly predict the minimal number of infected nodes for large recovery rates, either, induced by the fact that neighboring infectious states are highly covariant in sparse networks (see [52] and App. E.5), thus leading to false results in the times when the networks are sparse.

In order to accurately obtain the epidemic response curves and epidemic thresholds, a continuous-time non-absorbing simulation algorithm is necessary that should mimick the quasi-stationary state algorithms for static networks and discrete time. It is prematurely renamed to non-absorbing simulation algorithm here, since "quasi-stationary" becomes an odd name for substrates which are not stationary. Such an algorithm does not yet exist and should be the subject of future studies.

The epidemic thresholds evaluated in this section were merely qualitative evaluations of a critical control parameter. In more rigorous analyses, the epidemic threshold is defined by the value of the control parameter which maximizes the relative fluctuations of an order parameter. A continuous-time non-absorbing simulation algorithm could be used to correctly identify the relative fluctuations of both the endemic state as well as the minimal endemic state per period and thus settle the question whether the minimal endemic state should be used as the order parameter in temporal networks of varying activity or whether the average endemic state suffices.

Since rebinning data to larger time windows is a common approach to obtain the epidemic threshold, Gillespie simulations and the MCA were tested on rebinned networks. It was found that indeed, correspondence between both methods is increased on rebinned networks. Since rebinning drastically changes the single-

snapshot structure of temporal networks towards denser structures, covariance between neighboring infection states decreases and justifies the use of the MCA. After rebinning, however, the substrate is fundamentally different to the original structure and therefore predictions about the epidemic threshold are not predictions concerning the original system anymore.

Comparing SIS spreading results on Flockwork and edge activity model surrogate networks showed that solely mimicking single-snapshot structure and circadian activity is not sufficient to accurately predict the outcome of epidemic processes. Infected nodes appear to be changing too fast between disjoint neighborhoods, such that the epidemic response curves approach the mean-field limit. However, this random mixing is favourable for the MCA: on the surrogate networks, the approximation yields better results than on the real data. In real data, this kind of mixing does not occur, as the probability to reconnect to a recently disconnected neighbor is higher than to reconnect to other nodes. Furthermore, both the edge activity model as well as the Flockwork model surrogates yield response curves which are almost equal, despite having drastically different structures at times where the mean degree is large. Therefore, future studies should focus on the modeling and analysis of non-random mixing with memory, rather than single-snapshot structure.

9
Summary and Conclusions

The work presented in this dissertation analyzes basic spreading phenomena in static and temporal networks with a focus on human systems, aiming at advancing the understanding of how structural and temporal changes in those systems are connected to observed phenomena.

In Part II, the hypothesis was tested whether networks of modular hierarchical structure provide an optimal topology for fast diffusion and searches as modeled by random walks.

To this end, several existing generative network models were adapted such that, using a single control parameter, they interpolate between highly regular, locally clustered structures and the Erdős–Rényi random network model while keeping the mean degree fixed. A new generalized model formulation was presented which includes all presented explicit models and allows to generate small-world networks for arbitrary notions of distance. It was shown that the explicit models discussed here correspond to each other by (i) displaying the similarity of their network properties and (ii) comparing modular hierarchical network partitions an algorithm found by analyzing network samples of these models. All network models were found to generate structures which are interpretable to be modular hierarchical.

These models were used to evaluate the pair-averaged first passage time, an observable quantifying how fast randomly passed information can spread on these networks. For this purpose, a new heuristic was introduced to estimate this observable from network properties, for the first time including the clustering coefficient, such that networks do not have to be assumed to be locally tree-like for the heuristic to be applicable. It was found that network models whose node degree variance increases as a function of their structural parameter do indeed show a minimum in pair-averaged first passage time. Average medium approximations of these structures failed to explain the emergence of this effect, proving that network structure cannot be disregarded in this context. Another model which has constant degree variance, however, shows no minimum.

Similarly, the cover time shows a minimum for the network models with varying degree variance but no minimum for the remaining model, an effect confirmed by both simulations and a newly found

heuristic to estimate the cover time based on first passage time statistics.

Following these results, it has to be concluded that in general modular hierarchical networks do not possess an optimal structure for processes connected to random walks. This raises the question which other circumstances explain why those structures are so inherent to the real-world. A possible explanation is given as follows.

A considerable number of networks which are said to be of modular hierarchical structure could be networks whose nodes are actually embedded in some underlying space such that node pairs are associated with distances and consequently connected with a probability following a power-law decreasing with distance $P(r) \propto r^{-\kappa}$. This connection probability can be interpreted to be inversely proportional to the cost of establishing a connection to this distance. Since each long-range connection considerably decreases the time taken for information to spread, it is favorable to establish them. Depending on the nature of the system and the connections, however, the cost to establish such a long-distance edge can grow as a power-law $C(r) \propto r^{\mu}$. Assuming that edges are less likely to be built the more they cost, these networks would then resemble power-law small-world networks and be misinterpreted to be of pure modular hierarchical structure. Indeed, similar considerations have been made to explain the distance-dependent connection probability of airports via airlines [136].

In order to test this hypothesis, future research should focus on investigating real-world networks in detail for how they might have been generated and in what way the cost of establishing connections between node pairs can be determined. The inference of such wider classes of modular hierarchical community structure is an active area of research [28].

Furthermore, future research using models as defined in Ch. 4 to study their implications for dynamic systems should carefully investigate how their observables depend on the node degree variance since this might heavily influence their results.

In contrast to the static network descriptions investigated in Part II, more and more research has been focused on temporally resolved face-to-face contact networks for recent years. Consequently, there are many open questions to explore. In particular, the emergence of temporal and structural properties in empirically collected face-to-face contact networks and their consequences for epidemic spreading were investigated in Part III.

It was shown that the fragmentation of networks to small, densely connected groups observed in empirical data can be interpreted as an emergent property based on two node behaviors: Nodes can either disconnect to be isolated or dis- and reconnect to a new neighborhood. Both these behaviors are associated with activity rates. For the first time, this perspective allows to classify the mixing-time scale of a temporal network based on a node property similar to the mean degree of a static network: temporal mixing is quantified by the in-

troduction of a *contact-renewal number* which counts the times a node changes its neighborhood per unit time. In combination with the average mean degree, this allows for a strict classification of both network density and temporal mixing based on node properties.

From this perspective, it was shown that three empirically collected temporal networks follow similar trajectories within this state space of contact-renewal number and average mean degree, circadially varying between slow/sparse and fast/denser states. For the first time, it was possible to produce surrogate networks with corresponding activity trajectories based on the introduced node behaviors with a new temporal network model called Flockwork. To this end, an inference method was developed to measure the temporal change of the previously described basic node behavioral rates and consequently simulate surrogate networks in continuous time. It was argued that heterogeneity in the group-size distribution is most likely a consequence of the methodical practice to bin continuous-time contact data to intervals of discrete time. These findings highlight the value of continuous-time models where the dimension of time can be directly mapped to the temporal dimension of real-world data, such that model samples can be analyzed in the same way as real-world data. To the best knowledge of the author, the discussion presented here is one of the first instances where continuous-time temporal networks were analyzed as opposed to formulations in discrete time—at least in the context of face-to-face contact networks. While some network properties could be explained within this picture, other observations of real data, such as heterogeneously distributed (inter-)contact and group life times could not be reproduced. It was argued that the introduction of node-specific interaction rates and connection probabilities based on an underlying social network could be used to increase the realism of the models regarding these observables. On this account, future studies could explore a hybrid modular hierarchical/Flockwork model where temporal mixing follows the previously mentioned node behaviors, but connections are only allowed to be established according to social distance, following the logics of the small-world network models discussed in Part II. This would allow for a systematic investigation of the influence of temporal link heterogeneity on network properties and on dynamic systems. In any case, the model introduced in Ch. 7 is a valuable contribution to the zoo of temporal network models due to its simplicity and accessibility, and insights it already provided here.

In an initial application of the introduced model, epidemic spreading was investigated on temporal face-to-face contact networks by means of the SIS model. In particular, the influence of group structure and circadian activity on the epidemic threshold were previously unexplored. To this end, the newly introduced classification of temporal networks by average mean degree and contact-renewal number was helpful to systematically investigate the influence of structural differences and differences in time scales of both temporal networks and the spreading diseases. To the best knowledge of

the author, this is the first time that *continuous-time* temporal networks were investigated as a substrate for a *continuous-time* epidemic spreading process.

Concerning the influence of group structure it was found that a temporal network topology of homogenous group-size distributions increases the epidemic threshold. Heterogeneous group-size distributions are associated with a decreased epidemic threshold—if the time-scale of the disease becomes comparable to the time-scale of the structural change. Furthermore, it was found that sparse and slow networks of random structure can be associated with an increased epidemic threshold, as compared to their static, aggregated counterparts, rendering temporally resolved networks to be less susceptible to an epidemic. This result is in direct conflict with the results of other studies, which predict a decreased epidemic threshold, based on the application of the individual-based Markov-chain approximation and the simulation of reactive processes in discrete time instead of the actual Poisson process. In order to underline the validity of the continuous-time simulation results, it was shown analytically that even in the edge-case of an infinitely infectious disease there exist regions of the contact-renewal number and average mean degree in which the system becomes disease-free. These results suggest that (i) evaluating epidemic response curves using the individual-based Markov-chain approximation in discrete time and (ii) simulating the reactive process in discrete time are both methods which should be used with care as they may provide predictions diametrical to those of the correct continuous-time perspective.

This was further confirmed by studying continuous-time spreading processes on circadially varying temporal networks. One of the key results in this regard is the appearance of a jump in the epidemic response curve for fast diseases (i.e. short recovery times) in systems of finite size. Since the studied temporal networks follow trajectories in the previously mentioned state space they can reach values of mean degree and contact-renewal number which would lead to a disease-free system. If the system remains in this region long enough, the disease-free state can be approached. The temporally-averaged number of infected people in the dynamic equilibrium is usually used as an order parameter to find the epidemic threshold. In contrast, this investigation suggests that the minimum number of infected people approached on the state-space trajectory should be used as the order parameter. It was shown that the individual-based Markov-chain approximation, on which previous studies were based, may yield strong differences in this observable, therefore potentially drastically underestimating the epidemic threshold. Apart from the usage of this approximation, studies further binned temporal networks to time intervals of lengths which were magnitudes larger than the original resolution. This practice increases the correspondence of simulation and Markov-chain approximation but also drastically changes the actual structure of the temporal networks such that results regarding these binned structures have to be met

with skepticism.

Studying circadially varying surrogate networks of (i) random structure and (ii) fragmented group structure it was found that their structural differences did not significantly change the epidemic response curve, at least for the disease time-scales studied here. One may expect to find differences between the structures for even faster diseases.

The results presented here highlight some methodological difficulties of research concerning temporally varying networks. Often, the use of discrete time only approximates both the original structure as well as the original processes. If the time length of a discrete time step is chosen too large, the results might differ drastically. Discrete-time descriptions are usually used to simplify evaluations and simulations. They lose their validity, however, when their results suggest conclusions diametrical to the results of the original, continuous-time processes. Therefore, future research should focus on new theoretical and numerical techniques to efficiently and correctly analyze spreading processes on temporal networks.

To sum up, the work presented in this dissertation advances the understanding of how certain structural and temporal properties of human systems influence *basic* spreading processes. Previously unknown relationships between the structural influence on random message passing were found and model networks were generalized to ease analyses. It was shown that several often-used approximations are not valid in certain circumstances even for those basic spreading processes discussed here. While much research already concentrates on more complex spreading phenomena, e.g. the feedback between structural change and dynamic evolution, this work shows the value of accurately understanding the consequences of simple systems, before flawed methods are applied to even more complex situations.

10
Acknowledgments

Without the help and support of a myriad of people and organizations this thesis would not have come to life. These pages are for them.

First, I would like to acknowledge the organizational support by the Humboldt University of Berlin and the Humboldt Graduate School in particular, whose 'Research Track' program greatly eased my transition from a newly minted M.Sc. to a doctoral candidate. As I spent most of my working days at their offices, I would like to thank the Robert Koch Institute for their financial support, taking the risk to fund a highly inter-disciplinary research project, and the possibility to travel to so many different places to present our research and to learn new things. For both financial as well as organizational help, I would like to express my gratitude towards the Integrated Research Institute for Life Sciences at HU Berlin as well as the Joachim Herz Stiftung, whose 'Add-on Fellowship for Interdisciplinary Life Science' provided valuable funds and networking possibilities towards the end of my journey.

Regarding personal support, I first want to thank Prof. Dr. Dirk Brockmann, who was not only a supervisor but an academic and personal mentor. It is hard to only take a few words to express how much support I have received by him, regarding our research, funding opportunities, traveling possibilities, guidance towards building an academic career, and for the continued provision of crêpes funding. Second, I would like to thank Prof. Dr. Igor Sokolov, with whom my endeavor of becoming a physicist began in the Fall of 2008 when I sat in my first physics lecture and both eagerly and confusedly looked at what appeared to be theoretical physics he'd scribbled on the large black board in the Gerthsen lecture hall. I am deeply thankful for him taking my highly inter-disciplinary dissertation project under his wings, such that the plane ride I started in 2008 is now guided to the runway by the same tower. I would furthermore like to thank Prof. Dr. Sune Lehmann for his general support during the PhD, help with his group's data, fruitful discussions, and agreeing to review the thesis without hesitation.

Research projects rarely come to a successful end without the aid of collaboration. Indeed, I had the opportunity to work together with a handful of people without whom my findings would have

been obfuscated or simply remain hidden. I would first like to thank Dr. Cristián Huepe, whose ability to take a step back and evaluate the big picture greatly improved our manuscript and showed me that some of the paths I had taken were actually dead ends. Further, I am thankful to Ulf Aslak for his collaboration with the very fun network visualization project "Netwulf", which I hope will be of great benefit to the network science community in the future. Lastly, I would like to thank George C. Cantwell and Alice Schwarze on behalf of our whole SFI summer school team for a fun collaboration unrelated to my thesis that nevertheless taught me things I was able to apply to my main projects.

Even though some might argue otherwise, working on a dissertation in theoretical physics is not *always* the exciting roller coaster ride one might think—instead, one sits in an office all day, riding within their own mind. Without the presence and support of colleagues one might not find this ride's exit and instead keep driving (themselves insane). Therefore, I would first like to thank my colleague Dr. Olga Baranov, for the smiles she put on my face, the discussions we had, and the frustrations we shared. I truly miss our everyday banter. Second, I want to thank Frank Schlosser and Dr. Paola Carillo-Bustamante, for interesting discussions, shared sorrows, and the highly appreciated editing work on parts of this thesis. Also, many thanks to Thea Denell, for making me laugh and choosing my lunch when I was not able to make any decisions anymore. Of course, I had the opportunity to share offices with many more people whom I have to leave unmentioned here—you were all an integrate part of my life for varying amounts of time and for that I am grateful.

When thanking colleagues for support one should not forget that ideally, a large proportion of one's social life is not spent in an office but with friends and family. I would first like to thank Dr. Florian Klimm and Dr. Marc Wiedermann. Not only did they support me tremendously as friends, they edited large portions of this thesis. Without their suggestions, it would likely be a fuzzier hairball of information. Concerning editing I also really want to express my gratitude towards Dr. Andreas Koher, Dr. Sarah Griffin, Lorenz Drescher, and Susanne Klimroth, who spontaneously jumped in to help during the final stages. I would further like to thank my other friends Alex, Winnie, Jakob, Jonas, Anna, Marie, and a multitude of people I have to leave unmentioned, for being such strong emotional support over the years, and for being understanding of me disappearing from social activities in the last year. Yet, mentioning emotional support, Wuwu's role in my life should not be left unacknowledged. Thank you, little guy, for all the laughs.

Without my family, I would not have been able to pursue an academic profession. I would like to express my gratitude to my mother Carola and my fathers Berndt and Ralf, for being supportive of and non-intervening in my life choices and for raising me to become a human being with a curious mind. I further want to thank my uncle

Rolf and my aunt Monika for providing stability in turbulent times.

Even though dissertation-writing guides generally advise against it, I would like to take the opportunity to thank my therapist for showing me that even non-academic problems are better solved by not tackling them alone. She helped me during a particularly hard time of my life. Any graduate student reading these acknowledgments should know that they are not alone and that accepting professional help is not a sign of weakness.

Last but not least, all of this would not have been possible without my partner Susi. Not only do I love *you*, but also the relationship of trust, support, and frankly, the life we have built together in the last years. Thank you, for being you.

Bibliography

[1] Watts, D. J. and Strogatz, S. H. "Collective dynamics of 'small-world' networks". *Nature* 393.6684 (June 1998), pp. 440–442.

[2] Page, L., Brin, S., Motwani, R., and Winograd, T. *The PageRank Citation Ranking: Bringing Order to the Web*. Technical Report 1999-66. Stanford InfoLab, Nov. 1999.

[3] Brockmann, D. and Helbing, D. "The Hidden Geometry of Complex, Network-Driven Contagion Phenomena". *Science* 342.6164 (Dec. 2013), pp. 1337–1342.

[4] Iannelli, F., Koher, A., Brockmann, D., Hövel, P., and Sokolov, I. M. "Effective distances for epidemics spreading on complex networks". *Physical Review E* 95.1 (Jan. 2017), p. 012313.

[5] Sekara, V., Stopczynski, A., and Lehmann, S. "The fundamental structures of dynamic social networks". *arXiv:1506.04704 [physics]* (June 2015).

[6] Maier, B. F. *B. F. Maier's FB friends network*. Repository, github.com/benmaier/BFMaierFBnetwork. Apr. 2017.

[7] Hui Tian, Hong Shen, and Matsuzawa, T. "RandomWalk Routing for Wireless Sensor Networks". *Sixth International Conference on Parallel and Distributed Computing Applications and Technologies (PDCAT'05)*. Dec. 2005, pp. 196–200.

[8] Travers, J. and Milgram, S. "An Experimental Study of the Small World Problem". *Sociometry* 32.4 (1969), pp. 425–443.

[9] Barrat, A., Barthelemy, M., and Vespignani, A. *Dynamical processes on complex networks*. First paperback edition. Cambridge: Cambridge University Press, 2013.

[10] Wu, F. Y. "Theory of resistor networks: the two-point resistance". *Journal of Physics A: Mathematical and General* 37.26 (2004), p. 6653.

[11] Christakis, N. A. and Fowler, J. H. "Social Network Sensors for Early Detection of Contagious Outbreaks". *PLOS ONE* 5.9 (Sept. 2010), e12948.

[12] Lusseau, D., Schneider, K., Boisseau, O. J., Haase, P., Slooten, E., and Dawson, S. M. "The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations". *Behavioral Ecology and Sociobiology* 54.4 (Sept. 2003), pp. 396–405.

[13] Watts, D. J., Dodds, P. S., and Newman, M. E. J. "Identity and Search in Social Networks". *Science* 296.5571 (May 2002), pp. 1302–1305.

[14] Clauset, A., Moore, C., and Newman, M. E. J. "Hierarchical structure and the prediction of missing links in networks". *Nature* 453.7191 (2008), pp. 98–101.

[15] Meunier, D., Lambiotte, R., Fornito, A., Ersche, K. D., and Bullmore, E. T. "Hierarchical modularity in human brain functional networks". *Frontiers in Neuroinformatics* 3 (Oct. 2009).

[16] Meunier, D., Lambiotte, R., and Bullmore, E. T. "Modular and hierarchically modular organization of brain networks". *Frontiers in Neuroscience* 4 (2010), p. 200.

[17] Kaiser, M. and Simonotto, J. "Limited spreading: How hierarchical networks prevent the transition to the epileptic state". *Modeling Phase Transitions in the Brain*. Ed. by Steyn-Ross, D. A. and Steyn-Ross, M. Springer Series in Computational Neuroscience 4. Springer New York, 2010, pp. 99–116.

[18] Robinson, P. A., Henderson, J. A., Matar, E., Riley, P., and Gray, R. T. "Dynamical reconnection and stability constraints on cortical network architecture". *Physical Review Letters* 103.10 (Sept. 2009), p. 108104.

[19] Sarkar, S., Henderson, J. A., and Robinson, P. A. "Spectral characterization of hierarchical network modularity and limits of modularity detection". *PLoS ONE* 8.1 (Jan. 2013), e54383.

[20] Klimm, F., Bassett, D. S., Carlson, J. M., and Mucha, P. J. "Resolving structural variability in network models and the brain". *PLOS Computational Biology* 10.3 (Mar. 2014), e1003491.

[21] Sporns, O. *Networks of the brain*. OCLC: ocn551342282. Cambridge, Mass: MIT Press, 2011.

[22] Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., and Barabási, A.-L. "Hierarchical Organization of Modularity in Metabolic Networks". *Science* 297.5586 (Aug. 2002), pp. 1551–1555.

[23] Barabási, A.-L. and Oltvai, Z. N. "Network biology: understanding the cell's functional organization". *Nature Reviews Genetics* 5.2 (Feb. 2004), pp. 101–113.

[24] Sales-Pardo, M., Guimerà, R., Moreira, A. A., and Amaral, L. A. N. "Extracting the hierarchical organization of complex systems". *Proceedings of the National Academy of Sciences of the United States of America* 104.39 (Sept. 2007), pp. 15224–15229.

[25] Yerra, B. M. and Levinson, D. M. "The emergence of hierarchy in transportation networks". *The Annals of Regional Science* 39.3 (Sept. 2005), pp. 541–553.

[26] Smith, C., Puzio, R. S., and Bergman, A. "Hierarchical network structure promotes dynamical robustness". *arXiv:1412.0709* (Dec. 2014).

[27] Webster, J. R. "Hierarchical organization of ecosystems". *Theoretical Systems Ecology*. Academic Press, 1979, pp. 119–129.

[28] Peixoto, T. P. "Hierarchical block structures and high-resolution model selection in large networks". *Physical Review X* 4.1 (2014), p. 011047.

[29] Rosvall, M. and Bergstrom, C. T. "Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems". *PLoS ONE* 6.4 (Apr. 2011), e18209.

[30] Pan, R. K. and Sinha, S. "Modular networks with hierarchical organization: The dynamical implications of complex structure". *Pramana* 71.2 (Aug. 2008), pp. 331–340.

[31] Kleinberg, J. "The Small-world Phenomenon: An Algorithmic Perspective". *Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing*. STOC '00. New York, NY, USA: ACM, 2000, pp. 163–170.

[32] Newman, M. E. J. *Networks*. Second edition. Oxford, United Kingdom ; New York, NY, United States of America: Oxford University Press, 2018.

[33] Klafter, J. and Sokolov, I. M. *First steps in random walks: from tools to applications*. OCLC: ocn714724924. Oxford ; New York: Oxford University Press, 2011.

[34] Berg, H. C. *Random walks in biology*. Princeton, NJ: Princeton University Press, 1993.

[35] Holme, P. and Saramäki, J. "Temporal networks". *Physics Reports*. Temporal Networks 519.3 (2012), pp. 97–125.

[36] Bansal, S., Read, J., Pourbohloul, B., and Meyers, L. A. "The dynamic nature of contact networks in infectious disease epidemiology". *Journal of Biological Dynamics* 4.5 (Sept. 2010), pp. 478–489.

[37] Xie, X., Li, Y., Chwang, A. T. Y., Ho, P. L., and Seto, W. H. "How far droplets can move in indoor environments–revisiting the Wells evaporation-falling curve". *Indoor Air* 17.3 (June 2007), pp. 211–225.

[38] Taubenberger, J. K. and Morens, D. M. "The Pathology of Influenza Virus Infections". *Annual Review of Pathology: Mechanisms of Disease* 3.1 (Feb. 2008), pp. 499–522.

[39] Stopczynski, A., Sekara, V., Sapiezynski, P., Cuttone, A., Madsen, M. M., Larsen, J. E., and Lehmann, S. "Measuring Large-Scale Social Networks with High Resolution". *PLOS ONE* 9.4 (Apr. 2014), e95978.

[40] Barrat, A., Cattuto, C., Tozzi, A. E., Vanhems, P., and Voirin, N. "Measuring contact patterns with wearable sensors: methods, data characteristics and applications to data-driven simulations of infectious diseases". *Clinical Microbiology and Infection* 20 (Jan. 2014), pp. 10–16.

[41] Mastrandrea, R., Fournet, J., and Barrat, A. "Contact Patterns in a High School: A Comparison between Data Collected Using Wearable Sensors, Contact Diaries and Friendship Surveys". *PLOS ONE* 10.9 (Jan. 2015), e0136497.

[42] Isella, L., Stehlé, J., Barrat, A., Cattuto, C., Pinton, J.-F., and Van den Broeck, W. "What's in a crowd? Analysis of face-to-face behavioral networks." *Journal of Theoretical Biology* 271 (Feb. 2011), pp. 166–180.

[43] Zhao, K., Stehlé, J., Bianconi, G., and Barrat, A. "Social network dynamics of face-to-face interactions". *Physical review. E, Statistical, nonlinear, and soft matter physics* 83 (2011).

[44] Vestergaard, C. L., Génois, M., and Barrat, A. "How memory generates heterogeneous dynamics in temporal networks". *Physical Review E* 90.4 (Oct. 2014), p. 042805.

[45] Speidel, L., Klemm, K., Eguíluz, V. M., and Masuda, N. "Temporal interactions facilitate endemicity in the susceptible-infected-susceptible epidemic model". *New Journal of Physics* 18.7 (July 2016), p. 073013.

[46] Perra, N., Gonçalves, B., Pastor-Satorras, R., and Vespignani, A. "Activity driven modeling of time varying networks". *Scientific Reports* 2 (June 2012), p. 469.

[47] Malmgren, R. D., Stouffer, D. B., Motter, A. E., and Amaral, L. A. N. "A Poissonian explanation for heavy tails in e-mail communication". *Proceedings of the National Academy of Sciences* 105.47 (Nov. 2008), pp. 18153–18158.

[48] Malmgren, R. D., Stouffer, D. B., Campanharo, A. S. L. O., and Amaral, L. A. N. "On Universality in Human Correspondence Activity". *Science* 325.5948 (Sept. 2009), pp. 1696–1700.

[49] Chakrabarti, D., Wang, Y., Wang, C., Leskovec, J., and Faloutsos, C. "Epidemic thresholds in real networks". *ACM Transactions on Information and System Security* 10.4 (Jan. 2008), pp. 1–26.

[50] Gómez, S., Arenas, A., Borge-Holthoefer, J., Meloni, S., and Moreno, Y. "Discrete-time Markov chain approach to contact-based disease spreading in complex networks". *EPL (Europhysics Letters)* 89.3 (Feb. 2010), p. 38009.

[51] Wang, Y., Chakrabarti, D., Wang, C., and Faloutsos, C. "Epidemic spreading in real networks: an eigenvalue viewpoint". *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.* 2003, pp. 25–34.

[52] Givan, O., Schwartz, N., Cygelberg, A., and Stone, L. "Predicting epidemic thresholds on complex networks: Limitations of mean-field approaches". *Journal of Theoretical Biology* 288 (Nov. 2011), pp. 21–28.

[53] Anonymous. "Influenza in a boarding school". *British Medical Journal* 4 (Mar. 1978), p. 587.

[54] Brauer, F., Van den Driessche, P., Wu, J., and Allen, L. J. S., eds. *Mathematical epidemiology*. Mathematical biosciences subseries 1945. OCLC: ocn229446076. Berlin: Springer, 2008.

[55] Keeling, M. J. and Rohani, P. *Modeling infectious diseases in humans and animals*. OCLC: 751695215. Princeton: Princeton University Press, 2011.

[56] Valdano, E., Ferreri, L., Poletto, C., and Colizza, V. "Analytical Computation of the Epidemic Threshold on Temporal Networks". *Physical Review X* 5.2 (Apr. 2015), p. 021005.

[57] Valdano, E., Fiorentin, M. R., Poletto, C., and Colizza, V. "Epidemic Threshold in Continuous-Time Evolving Networks". *Physical Review Letters* 120.6 (Feb. 2018).

[58] Hermanussen, M., Danker-Hopfe, H., and Weber, G. W. "Body weight and the shape of the natural distribution of weight, in very large samples of German, Austrian and Norwegian conscripts". *International Journal of Obesity* 25.10 (Oct. 2001), p. 1550.

[59] Reed, W. J. and Jorgensen, M. "The Double Pareto-Lognormal Distribution—A New Parametric Model for Size Distributions". *Communications in Statistics - Theory and Methods* 33.8 (Dec. 2004), pp. 1733–1753.

[60] Dijkstra, E. W. "A note on two problems in connexion with graphs". *Numerische Mathematik* 1.1 (Dec. 1959), pp. 269–271.

[61] Fortunato, S. "Community detection in graphs". *Physics Reports* 486.3–5 (Feb. 2010), pp. 75–174.

[62] Lancichinetti, A., Fortunato, S., and Kertesz, J. "Detecting the overlapping and hierarchical community structure of complex networks". *New Journal of Physics* 11.3 (Mar. 2009), p. 033015.

[63] Albert, R. and Barabasi, A.-L. "Statistical mechanics of complex networks". *Reviews of Modern Physics* 74.1 (Jan. 2002). arXiv: cond-mat/0106096, pp. 47–97.

[64] Newman, M. E. J. "Modularity and community structure in networks". *Proceedings of the National Academy of Sciences* 103.23 (June 2006), pp. 8577–8582.

[65] Simon, H. A. "The architecture of complexity". *Proceedings of the American Philosophical Society* 106.6 (1962), pp. 467–482.

[66] Arenas, A., Díaz-Guilera, A., and Pérez-Vicente, C. J. "Synchronization reveals topological scales in complex networks". *Physical Review Letters* 96.11 (2006), p. 114102.

[67] Rosvall, M. and Bergstrom, C. T. "Maps of random walks on complex networks reveal community structure". *Proceedings of the National Academy of Sciences* 105.4 (Jan. 2008), pp. 1118–1123.

[68] Rosvall, M., Axelsson, D., and Bergstrom, C. T. "The map equation". *The European Physical Journal Special Topics* 178.1 (Nov. 2009), pp. 13–23.

[69] Fire, M., Katz, G., Elovici, Y., Shapira, B., and Rokach, L. "Predicting Student Exam's Scores by Analyzing Social Network Data". *Active Media Technology*. Ed. by Huang, R., Ghorbani, A. A., Pasi, G., Yamaguchi, T., Yen, N. Y., and Jin, B. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2012, pp. 584–595.

[70] Krebs, V. "Mapping networks of terrorist cells". *Connections* 24 (2002), pp. 43–52.

[71] Ahn, Y.-Y., Bagrow, J. P., and Lehmann, S. "Link communities reveal multiscale complexity in networks". *Nature* 466.7307 (Aug. 2010), pp. 761–764.

[72] Kim, J. H. and Vu, V. H. "Generating random regular graphs". *Proceedings of the thirty-fifth ACM symposium on Theory of computing - STOC '03*. San Diego, CA, USA: ACM Press, 2003, p. 213.

[73] Gillespie, D. T. "Exact stochastic simulation of coupled chemical reactions". *The Journal of Physical Chemistry* 81.25 (1977), pp. 2340–2361.

[74] Batagelj, V. and Brandes, U. "Efficient generation of large random networks". *Physical Review E* 71.3 (2005), p. 036113.

[75] Gilbert, E. N. "Random Graphs". *The Annals of Mathematical Statistics* 30.4 (Dec. 1959), pp. 1141–1144.

[76] Erdős, P. and Rényi, A. "On random graphs I". *Publicationes Mathematicae (Debrecen)* 6 (1959), pp. 290–297.

[77] Dunbar, R. "Neocortex size as a constraint on group size in primates". *Journal of Human Evolution* 22.6 (June 1992), pp. 469–493.

[78] Dall, J. and Christensen, M. "Random geometric graphs". *Physical Review E* 66.1 (July 2002), p. 016121.

[79] Sekara, V. and Lehmann, S. "The Strength of Friendship Ties in Proximity Sensor Data". *PLOS ONE* 9.7 (July 2014), e100915.

[80] Paul W. Holland, K. B. L. "Stochastic blockmodels: First steps". *Social Networks* 5.2 (1983), pp. 109–137.

[81] Karrer, B. and Newman, M. E. J. "Stochastic blockmodels and community structure in networks". *Physical Review E* 83.1 (Jan. 2011), p. 016107.

[82] Barabási, A.-L. and Albert, R. "Emergence of scaling in random networks". *Science* 286.5439 (Oct. 1999), pp. 509–512.

[83] Bollobás, B., Riordan, O., Spencer, J., and Tusnády, G. "The degree sequence of a scale-free random graph process: Degree Sequence of a Random Graph". *Random Structures & Algorithms* 18.3 (May 2001), pp. 279–290.

[84] Holme, P. "Modern temporal network theory: a colloquium". *The European Physical Journal B* 88.9 (Sept. 2015), p. 234.

[85] Zachary, W. W. "An information flow model for conflict and fission in small groups". *Journal of Anthropological Research* 33 (1977), pp. 452–473.

[86] Aslak, U., Rosvall, M., and Lehmann, S. "Constrained information flows in temporal networks reveal intermittent communities". *Physical Review E* 97.6 (June 2018), p. 062312.

[87] Koher, A., Lentz, H. H. K., Hövel, P., and Sokolov, I. M. "Infections on Temporal Networks—A Matrix-Based Approach". *PLOS ONE* 11.4 (Jan. 2016), e0151209.

[88] Cattuto, C., Broeck, W. V. d., Barrat, A., Colizza, V., Pinton, J.-F., and Vespignani, A. "Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks". *PLOS ONE* 5.7 (July 2010), e11596.

[89] Vestergaard, C. L. and Génois, M. "Temporal Gillespie Algorithm: Fast Simulation of Contagion Processes on Time-Varying Networks". *PLOS Computational Biology* 11.10 (Oct. 2015), e1004579.

[90] Zhang, X., Moore, C., and Newman, M. E. J. "Random graph models for dynamic networks". *The European Physical Journal B* 90.10 (Oct. 2017), p. 200.

[91] Valdano, E., Poletto, C., and Colizza, V. "Infection propagator approach to compute epidemic thresholds on temporal networks: impact of immunity and of limited temporal resolution". *The European Physical Journal B* 88.12 (Dec. 2015), p. 341.

[92] Maier, B. F. and Brockmann, D. "Cover time for random walks on arbitrary complex networks". *Physical Review E* 96.4 (Oct. 2017), p. 042307.

[93] Maier, B. F., Huepe, C., and Brockmann, D. "Modular hierarchical and power-law small-world networks bear structural optima for minimal first passage times and cover time". *Journal of Complex Networks* (Apr. 2019).

[94] Maier, B. F. "Generalization of the small-world effect on a model approaching the Erdős–Rényi random graph". *Scientific Reports* 9.1 (June 2019), p. 9268.

[95] Øksendal, B. *Stochastic differential equations: An introduction with applications*. Universitext. OCLC: 246776666. Berlin: Springer, 1992.

[96] Masuda, N., Porter, M. A., and Lambiotte, R. "Random walks and diffusion on networks". *arXiv:1612.03281* (Dec. 2016).

[97] Brockmann, D. "Anomalous diffusion and the structure of human transportation networks". *The European Physical Journal Special Topics* 157.1 (Apr. 2008), pp. 173–189.

[98] Hufnagel, L., Brockmann, D., and Geisel, T. "Forecast and control of epidemics in a globalized world". *Proceedings of the National Academy of Sciences* 101.42 (Oct. 2004), pp. 15124–15129.

[99] Lovász, L. "Random walks on graphs: a survey". *Combinatorics, Paul Erdős is Eighty*. Ed. by Miklós, D., Sós, V. T., and Szőnyi, T. Vol. 2. Budapest: János Bolyai Mathematical Society, 1996, pp. 353–398.

[100] Lau, H. W. and Szeto, K. Y. "Asymptotic analysis of first passage time in complex networks". *EPL (Europhysics Letters)* 90.4 (2010), p. 40005.

[101] Lin, Y., Julaiti, A., and Zhang, Z. "Mean first-passage time for random walks in general graphs with a deep trap". *The Journal of Chemical Physics* 137.12 (Sept. 2012), p. 124104.

[102] Melnik, S., Hackett, A., Porter, M. A., Mucha, P. J., and Gleeson, J. P. "The unreasonable effectiveness of tree-based theory for networks with clustering". *Physical Review E* 83.3 (Mar. 2011), p. 036112.

[103] Sood, V., Redner, S., and ben-Avraham, D. "First-passage properties of the Erdős–Renyi random graph". *Journal of Physics A: Mathematical and General* 38.1 (2005), p. 109.

[104] Thiel, F. and Sokolov, I. M. "Effective-medium approximation for lattice random walks with long-range jumps". *Physical Review E* 94.1 (July 2016), p. 012135.

[105] Tavani, F. and Agliari, E. "First-passage phenomena in hierarchical networks". *Physical Review E* 93.2 (Feb. 2016), p. 022133.

[106] Domenico, M. D., Solé-Ribalta, A., Gómez, S., and Arenas, A. "Navigability of interconnected networks under random failures". *Proceedings of the National Academy of Sciences* 111.23 (June 2014), pp. 8351–8356.

[107] Abdullah, M. "The cover time of random walks on graphs". PhD thesis. Feb. 2012.

[108] Cardanobile, S. "Diffusion systems and heat equations on networks". *arXiv:0807.2362* (July 2008).

[109] Mohar, B. "Some applications of Laplace eigenvalues of graphs". *Graph Symmetry*. NATO ASI Series. Springer, Dordrecht, 1997, pp. 225–275.

[110] Tetali, P. "Random walks and the effective resistance of networks". *Journal of Theoretical Probability* 4.1 (Jan. 1991), pp. 101–109.

[111] Cooper, C. and Frieze, A. "The cover time of sparse random graphs". *Random Structures and Algorithms* 30.1-2 (Jan. 2007), pp. 1–16.

[112] Cooper, C. and Frieze, A. "The cover time of the preferential attachment graph". *Journal of Combinatorial Theory, Series B* 97.2 (2007), pp. 269–290.

[113] Kermack, W. O. and McKendrick, A. G. "Contributions to the mathematical theory of epidemics—I". *Bulletin of Mathematical Biology* 53.1-2 (Mar. 1991), pp. 33–55.

[114] Vestergaard, C. L., Valdano, E., Génois, M., Poletto, C., Colizza, V., and Barrat, A. "Impact of spatially constrained sampling of temporal contact networks on the evaluation of the epidemic risk". *European Journal of Applied Mathematics* 27.6 (Dec. 2016), pp. 941–957.

[115] Van Mieghem, P. *Performance analysis of complex networks and systems*. Cambridge: Cambridge University Press, 2014.

[116] Chung, F., Lu, L., and Vu, V. "Spectra of random graphs with given expected degrees". *Proceedings of the National Academy of Sciences* 100.11 (May 2003), pp. 6313–6318.

[117] Oliveira, M. M. de and Dickman, R. "How to simulate the quasistationary state". *Physical Review E* 71.1 (Jan. 2005), p. 016129.

[118] Maier, B. F. *cMHRN - A C++/Python/MATLAB package to generate SSMH and PLSW networks in a fast manner*. Repository, github.com/benmaier/cMHRN. Mar. 2018.

[119] Song, H. F. and Wang, X.-J. "Simple, distance-dependent formulation of the Watts-Strogatz model for directed and undirected small-world networks". *Physical Review E* 90.6 (Dec. 2014), p. 062801.

[120] Maier, B. F. *smallworld - A Python package for generating and analyzing networks from the alternative small-world model*. Repository, github.com/benmaier/smallworld. Nov. 2018.

[121] Watts, D. J. "How small is the world, really?" *Medium* (Feb. 2016).

[122] Kleinfield, J. "Six Degrees: Urban Myth?" *Psychology Today* (Mar. 2002).

[123] Cross, R. and Parker, A. *The hidden power of social networks*. Boston, MA: Harvard Business School Press, 2004.

[124] Aicher, C., Jacobs, A. Z., and Clauset, A. "Learning latent block structure in weighted networks". *Journal of Complex Networks* 3.2 (June 2015), pp. 221–248.

[125] Colizza, V., Pastor-Satorras, R., and Vespignani, A. "Reaction-diffusion processes and metapopulation models in heterogeneous networks". *Nature Physics* 3.4 (Apr. 2007), pp. 276–282.

[126] Shen-Orr, S. S., Milo, R., Mangan, S., and Alon, U. "Network motifs in the transcriptional regulation network of Escherichia coli". *Nature Genetics* 31.1 (2002), pp. 64–68.

[127] Girvan, M. and Newman, M. E. J. "Community structure in social and biological networks". *Proceedings of the National Academy of Sciences of the United States of America* 99.12 (June 2002), pp. 7821–7826.

[128] Dawah, H. A., Hawkins, B. A., and Claridge, M. F. "Structure of the Parasitoid Communities of Grass-Feeding Chalcid Wasps". *Journal of Animal Ecology* 64.6 (1995), pp. 708–720.

[129] Knuth, D. E. *The Stanford GraphBase: A platform for combinatorial computing*. Reading, MA: Addison-Wesley, 1993.

[130] Steger, A. and Wormald, N. C. "Generating random regular graphs quickly". *Comb. Probab. Comput.* 8.4 (July 1999), pp. 377–396.

[131] Roth, C., Kang, S. M., Batty, M., and Barthelemy, M. "A long-time limit for world subway networks". *Journal of The Royal Society Interface* (2012).

[132] Elsässer, R. and Sauerwald, T. "Tight bounds for the cover time of multiple random walks". *Theoretical Computer Science*. Selected Papers from 36th International Colloquium on Automata, Languages and Programming (ICALP 2009) 412.24 (May 2011), pp. 2623–2641.

[133] Atkinson, K. E. *An introduction to numerical analysis*. 2nd. New York: Wiley, 1989.

[134] Halliday, D., Resnick, R., Walker, J., and Koch, S. W. *Physik*. 1. korr. Nachdr. OCLC: 297583264. Weinheim: Wiley-VCH, 2005.

[135] Diaz, J., Mitsche, D., and Perez, X. "Dynamic Random Geometric Graphs". *arXiv:cs/0702074* (Feb. 2007).

[136] Bianconi, G., Pin, P., and Marsili, M. "Assessing the relevance of node features for network structure". *Proceedings of the National Academy of Sciences* 106.28 (July 2009), pp. 11433–11438.

[137] Ravasz, E. and Barabási, A.-L. "Hierarchical organization in complex networks". *Physical Review E* 67.2 (Feb. 2003), p. 026112.

[138] Lambiotte, R. *Hierarchical networks: (information) diffusion and mapping*. Presentation. 2010.

[139] Philip, J. "The Probability Distribution of the Distance between two Random Points in a Box". Department of Mathematics, KTH, Dec. 2007.

[140] Ross, S. M. *Stochastic processes*. 2nd ed. Wiley series in probability and statistics. New York: Wiley, 1996.

[141] Andrae, R., Schulze-Hartung, T., and Melchior, P. "Dos and don'ts of reduced chi-squared". *arXiv:1012.3754* (Dec. 2010).

[142] Van Mieghem, P. and Cator, E. "Epidemics in networks with nodal self-infection and the epidemic threshold". *Physical Review E* 86.1 (July 2012), p. 016116.

Appendix

A
Infomap: A Flow-based Algorithm to Find Modular Hierarchical Network Partitions

Real–world static networks are often claimed to be of *modular hierarchical structure* [14, 15, 18, 19, 22, 24, 28, 62, 66, 137, 138], where this term usually refers to the fact that the nodes of a network can be grouped to form modules based on some sort of similarity which is computed from their connections. Then, recursively, these modules may again be grouped together to form an even coarser grained structure and so forth. A single realization of such a modular hierarchical grouping scheme is called a *modular hierarchical network partition* (MHNP). While there exist descriptions allowing for *overlapping* modules where nodes and submodules may be part of multiple groups, these will not be of concern within this work since they do not add any particular value to the qualitative investigation of the consequence of modular hierarchical clustering on dynamic sytems.

Extraction algorithms to find "optimal" such partitions usually search the space of all partitions of a particular observed network to find a single one or multiple which minimize a certain cost function (or maximize a certain likelihood function). Some of these depend on the description of the network as generated by a stochastic block model and maximize the likelihood that a nested stochastic block model defined by a modular hierarchical network partition will generate the observed network while minimizing the information theoretic description length of the model to prevent overfitting [14, 28]. However, this thesis mainly aims at exploring the consequence of modular hierarchical clustering on the outcome of dynamic processes, more accurately on the outcome of random walk observables. To this end, it seems plausible to deploy a modular hierarchical partition algorithm which finds partitions based on the similarity of nodes concerning their role in random walks. One such algorithm is *Infomap*, which searches for partitions minizing the so–called *map equation* [29, 67, 68]. In the following, the underlying mechanism of *Infomap* is briefly outlined to give an intuition about its inner workings, while a complete description would go beyond the scope of this work due to it only playing a minor role overall.

At the base of the algorithm lies the trajectory of a single random

walker traversing through the network as described in Sec. 3.1.1. Giving each node a unique codename formed by the bits "1" and "0", the trajectory of the walker can be completely encoded by the succession of these codenames, called a *trajectory sentence* in the subsequent description.[1] *Infomap* then aims to find the minimal description length of this trajectory sentence by introducing a modular hierarchical network partition.

[1] The codenames are chosen *prefix-free* such that no delimiting symbols are necessary.

Following the principle of Huffman coding, nodes which are visited more often than others are given shorter names such that their overall space in the trajectory sentence is minimized. This implies that node names can become rather long with increasing number of nodes N leading to a long trajectory sentence of any ergodic random walk. The idea behind *Infomap* is that (short) codenames for single nodes can be reused when nodes are grouped into modules in which the random walker spends a considerable amount of time steps before it exits the module to enter another module. Then, short module names can be used in the trajectory sentence, each indicating that it has been entered by the random walker. The codenames following this entry symbol consiquently belong to the nodes of this module, even though their names can be duplicates of node names in other modules. Additionally, exit code names are used to indicate that the current module has been left.

Instead of actually simulating a random walker and subsequently encoding its trajectory, the description length of a trajectory of n steps can be lower bounded by $n\mathcal{L}(M)$ where $\mathcal{L}(M)$ is the lower bound average description length of a single step given that the coding is based on the modular hierarchical network partition M. The lower bound can be evaluated using Shannon's source coding theorem which states that when a series of entities i is drawn randomly, each with probability p_i, from an ensemble $\mathcal{R}$, then the average length of the codename of such a symbol cannot be less than the ensemble's information entropy which is given as

$$H(\mathcal{R}) = -\sum_{i=1}^{|\mathcal{R}|} p_i \log_2 p_i. \tag{A.1}$$

The trajectory sentence of the random walk can be approximated as a randomly generated sentence of codenames associated with their probability of occurence. For a two-level modular network partition, nodes are grouped to modules, but those modules are not further grouped. In such a partition the following events with corresponding codenames may occur.

1. Each module is associated with an entry symbol. When there are m such modules, there have to be m entry codes, each occuring with conditional probability $q_{i\curvearrowright} / \sum_{j=1}^{m} q_{j\curvearrowright}$ (the probability that module i is exited given that any module switching event happened). The probability $q_{i\curvearrowright}$ is proportional to the number of edges leading out of the module. This ensemble of events shall be called $\mathcal{Q}$ in the following.

2. Each visiting node in a module is associated with a visiting symbol. The conditional probability that node u of module i is chosen given that any event of module i is chosen is $p_u / [q_i + \sum_{u \in i} p_u]$. The visiting probability p_u of node u is chosen as the equilibrium probability to find a random walker on node u as given in Eq. (3.3). Furthermore, each module i is additionally associated with an exit symbol, where $q_{i\curvearrowright}$ is the probability that module i is exited, hence, the corresponding conditional probability that the exit happened given that an event of module i happened is $q_{i\curvearrowright} / [q_{i\curvearrowright} + \sum_{u \in i} p_u]$. This particular ensemble of module-i events shall be called $\mathcal{P}^{(i)}$.

Considering these events and their ensembles $\mathcal{Q}$ and $\mathcal{P}^{(i)}$, the weighted averaged length of a codename used within a single step is lower bounded by

$$\mathcal{L}(M) = \left(\sum_{i=1}^{m} q_{i\curvearrowright} \right) H(\mathcal{Q}) + \sum_{i=1}^{m} \left(q_{i\curvearrowright} + \sum_{u \in i} p_u \right) H(\mathcal{P}^{(i)})$$

which is called the *map equation* (here, *map* is a synonym for two-level modular network partition). This average single-step codename description length can be computed in a fast manner for any two-level partition M. In principle, any standard algorithm to minimize a cost function can be used to find an optimal partition. For a more detailed explanation of the map equation the reader is referred to [67, 68].

For modular hierarchical network partitions, the map equation takes a recursive form. Any modular hierarchical network partition M consists, at first, of m submaps M^i. For each of these submaps the map equation is valid in the same form, where now exit probabilities refer to the exit of coarse grained modules. A detailed explanation of the hierarchical map equation and an algorithm to find an optimal hierarchical map is given in [29]. In this thesis, the C++-implementation of the method described and implemented in [29] was used with default parameter choices.

B
Evaluations Concerning Small-World Models

B.1 SSMH Clustering Coefficient

Picking up at Eq. (4.10), the long range contributions will be evaluated by considering the following scenario first. The pair $(1, j)$ has hierarchical distance ℓ_1, whereas the pair $(1, u)$ has hierarchical distance $\ell_1 < \ell_2$. This means that nodes 1 and j share a subgroup in layer ℓ_1 and thus, if the third node u has distance ℓ_2 to node 1, so has node j distance ℓ_2 to node u. Since the problem is symmetrical, the case $\ell_1 > \ell_2$ will contribute the same amount and hence it suffices to look at $\ell_1 < \ell_2$. Now there are $B^{\ell_1}(B-1)$ possible target nodes for node 1 in layer ℓ_1 and $B^{\ell_2}(B-1)$ possible target nodes in layer ℓ_2. Consequently, the total number of potential pairs in layer combination ℓ_1, ℓ_2 is $B^{\ell_1}(B-1) \times B^{\ell_2}(B-1)$ and the total sum evaluates to

$$\begin{aligned}\pi_L^{(3)} &= \sum_{\ell_1=1}^{L-1} \sum_{\ell_2=\ell_1+1}^{L} (B-1)^2 B^{\ell_1} B^{\ell_2} P_{\ell_1} P_{\ell_2} P_{\ell_2} \\ &= \frac{\langle k \rangle^3}{B-1} \left(\frac{1-\xi}{1-\xi^L} \right)^3 \frac{1}{1-\xi^2/B} \times \\ &\quad \times \left[\left(\frac{\xi^2}{B} \right) \frac{1-(\xi^3/B)^{L-1}}{1-\xi^3/B} - \left(\frac{\xi^2}{B} \right)^L \left(\frac{1-\xi^{L-1}}{1-\xi} \right) \right].\end{aligned}$$

The short-range contributions stem from pairs of target nodes (j, u) where node 1 has the same distance ℓ to both of them, hence they build a triad with probability P_ℓ^2. A distinction between two cases can be made.

1. Both j and u have distance ℓ to node 1, and are part of the same subgroup. This means that their distance is $\ell' < \ell$ and that they are connected with probability $P_{\ell'}$. The total number of possible pairs of distance $\ell' < \ell$ is $(1/2) \times B^{\ell}(B-1) \times B^{\ell'-1}(B-1)$. The additional factor $1/2$ emerges to avoid double counting (with once j as source and once u as source).

2. Both j and u have distance ℓ to node 1, but are not part of the same subgroup in $\ell' \leq \ell$. This means that u is at distance ℓ of

node j but the number of submodules that j can choose a neighbor from is reduced by two (its own subgroup and the subgroup of node 1). Hence, the total number of distinct pairs of this type is $(1/2) \times B^{\ell-1}(B-2)B^{\ell-1}(B-1)$. There are $B^{\ell-1}(B-1)$ nodes to pick as first neighbor of 1 and $B^{\ell-1}(B-2)$ nodes to pick as second neighbor of 1. Again, there is an additional factor $1/2$ to avoid double counting.

Considering these cases, the short-range contribution are as follows (where the Kronecker symbol δ_{ij} is used),

$$\begin{aligned}\pi_S^{(3)} &= \frac{1}{2}\sum_{\ell=1}^{L} P_\ell^2 \sum_{\ell'=1}^{\ell} B^{\ell-1}(B-1)B^{\ell'-1}(B-1-\delta_{\ell\ell'})P_{\ell'} \\ &= \frac{1}{2}\frac{\langle k\rangle^3}{(B-1)^2}\left(\frac{1-\xi}{1-\xi^L}\right)^3 (B-2)\frac{1-(\xi^3/B)^L}{1-\xi^3/B} + \\ &\quad + \frac{B-1}{1-\xi}\left(\frac{1-(\xi^2/B)^L}{1-\xi^2/B} - \frac{1-(\xi^3/B)^L}{1-\xi^3/B}\right).\end{aligned}$$

The contributions $\pi_L^{(2)}$ and $\pi_S^{(2)}$ can be calculated with an analogous procedure as above by setting $P_{\ell_2} = 1$ and $P_{\ell'} = 1$, respectively, which yields

$$\pi_L^{(2)} = \frac{\langle k\rangle^2(1-\xi)}{(1-\xi^L)^3}\left[\xi\frac{1-(\xi^2)^{L-1}}{1-\xi^2} - \frac{\xi^L-\xi^{2L-1}}{1-\xi}\right], \text{ and}$$
$$\pi_S^{(2)} = \frac{1}{2}\frac{\langle k\rangle^2}{B-1}\left(\frac{1-\xi}{1-\xi^L}\right)^2\left[(B-1)\frac{1-\xi^{2L}}{1-\xi^2} - \frac{1-(\xi^2/B)^L}{1-\xi^2/B}\right].$$

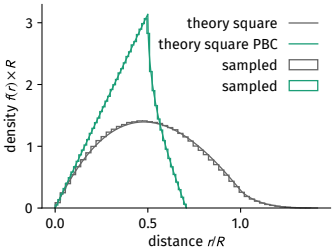


Figure B.1: Pairwise distance distributions of points uniformly distributed in a square $[0,R)^2$ with and without periodic boundary conditions. Sampling was done for $N = 1000$ points.

B.2 PLSW Networks

Consider the unit square $[0,1)^2$ with periodic boundary conditions (i.e. a 2-torus). On this space N nodes are distributed with uniform probability (i.e. their positions x and y follow the same uniform random distribution $x, y \sim \mathcal{U}(0,1)$). Since the torus is translationally invariant, each node can be assumed to be centered. Hence it suffices to consider a single node and compute the density of nodes around it at distance r. The problem is radially symmetric until radius $r_1 = 1/2$. Then, for $r > r_1$, each remaining corner of the square around the central node is congruent. Hence, the distance density $f(r)$ can be found by considering the split integral

$$1 = \int_0^{1/\sqrt{2}} dr\, f(r) = 4\int_0^{1/2} dr \int_0^{\pi/2} d\theta\, r + 4\int_{1/2}^{1/\sqrt{2}} dr \int_{\arccos(1/2r)}^{\pi/2-\arccos(1/2r)} d\theta\, r$$

which yields

$$f(r) = \begin{cases} 2\pi r, & \text{if } 0 \le r \le \frac{1}{2} \\ 2\pi r - 8r\arccos\left(\frac{1}{2r}\right), & \text{if } \frac{1}{2} < r \le \frac{1}{\sqrt{2}}. \end{cases} \tag{B.1}$$

This distribution is shown in Fig. B.1.

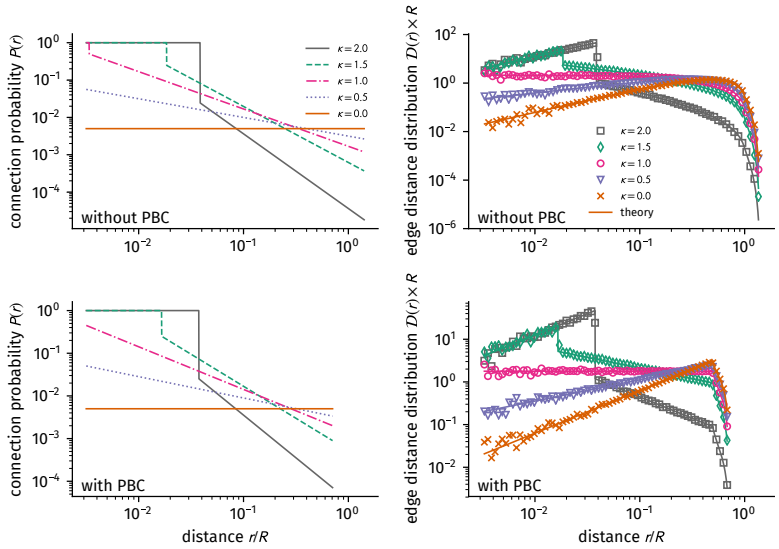


Figure B.2: Connection probability for random points in a square of side length R **(top row)** with and **(bottom row)** without periodic boundary conditions. **(Left column)** power-law connection probability $P(r)$ as defined by Eqs. (4.22) and (4.23), **(right column)** edge distance distribution $\mathcal{D}(r)$ (see Eq. (4.24)) with distance distributions **(bottom row)** Eq. (B.1) and **(top row)** Eq. (B.2). Parameters of choice were $N = 2000$, $\langle k \rangle = 10$, $\epsilon = 10^{-9}$, and $\kappa \in \{2, 1.5, 1, 0.5, 0\}$. For each parameter combination 20 networks were sampled.

When no periodic boundary conditions are considered, the distance distribution is

$$f(r) = 2rg(r^2) \tag{B.2}$$

with

$$g(s) = \begin{cases} -4\sqrt{s} + \pi + s & 0 < s \leq 1, \\ -2 - \pi - s + 4\arcsin(s^{-1/2}) + 4\sqrt{s-1} & 1 < s \leq 2, \end{cases}$$

as derived in [139].

As a numerical validation of the considerations in Sec. 4.3.1, networks were generated using the parameters $N = 2000$, $\langle k \rangle = 10$, $\epsilon = 10^{-9}$, and $\alpha \in \{2, 1.5, 1, 0.5, 0\}$ and the corresponding connection probability functions as defined by Eqs. (4.22) and (4.23), given the distance distributions Eqs. (B.1) and (B.2), respectively. The sampling algorithm was simply as follows. Start with an empty network of N nodes where each node u has coordinates $0 \leq x_u < 1$ and $0 \leq y_u < 1$, both drawn from a uniform distribution $\mathcal{U}(0,1)$. For each pair of nodes (u, v) add the edge (u, v) to the network with probability $P(r_{uv})$ where $r_{uv} = \sqrt{(x_u - x_v)^2 + (y_u - y_v)^2}$ for the simple unit square and $r_{uv} = \min\left\{\sqrt{(x_u - l_1 x_v)^2 + (y_u - l_2 y_v)^2} : (l_1, l_2) \in \{-1, 0, 1\}^2\right\}$ for the 2-torus. For each parameter combination 20 networks were generated, subsequently building a log-binned histogram to find the distribution of distances from the drawn edges. The numerically found edge distance distributions are indeed in close agreement[1] with Eq. (4.24), as shown in Fig. B.2.

[1] Small discrepancies are observed for short normalised distances r/R which is due to the small number of network samples and consequently small number of node pairs of short distances.

B.2.1 *Extension to a Finite Set of Distances*

The derivation of Sec. 4.3 is based on the assumption that the embedding topology is known. Assuming, instead, that the geometry of the embedding space is unknown but one has knowledge about a set of nodes and the distance of each pair of nodes, one may proceed as

follows. Let $\boldsymbol{r} \in (\mathbb{R}^+)^{N(N-1)/2}$ denote the increasingly sorted vector of all distances between pairs of nodes r_{uv}. Then, the normalization condition is

$$\sum_{e=1}^{N(N-1)/2} C_\kappa r_e^{-\kappa} = \frac{N \langle k \rangle}{2}$$

such that the prefactor C_κ evaluates to

$$C_\kappa = \frac{N \langle k \rangle}{2 \sum_{e=1}^{N(N-1)/2} r_e^{-\kappa}}.$$

Again, for small distances and small κ, the connection probability $P(r)$ will exceed unity. Hence, a minimum entry l of the vector $\boldsymbol{r}$ has to be chosen below which every distance will produce an edge while above this entry, edge creation will follow a power-law with exponent κ. This entry l is given as the last entry meeting the inequality

$$\sum_{e=1}^{l} \left[C_\kappa r_l^{-\kappa} - 1 \right] > 0.$$

Subsequently, edges can be drawn according to their index number in the vector $\boldsymbol{r}$

$$P_e = \begin{cases} 1 & e \leq l \\ C_\kappa r_e^{-\kappa} & e > l. \end{cases}$$

It is, however, of crucial importance to keep a mapping between the node pair (u, v) and the edge entry in the sorted vector $\boldsymbol{r}$. Furthermore, there will be an obvious bias when there are multiple edges of critical distance r_l. Note that if the pair-wise distances are sufficiently distinct from each other, l can be directly mapped to its distance r_l such that the connection probability may be written down as a dependence on r as

$$P(r) = \begin{cases} 1 & r \leq r_l \\ C_\kappa r^{-\kappa} & r > r_l. \end{cases}$$

At this point it might be important to note that r does not need to be a "distance" anymore in the sense that the values r_{uv} do not have to meet the triangle inequality. Instead, these values could stem from semimetrics but also represent rank values or other positive real numbers, i.e. general edge weights.

B.2.2 Extension to Categorical Distances

For the considerations above it was assumed that distances are positive real numbers and hence distributed according to a probability *density* function. However, there exist cases where distances are integer numbers (e.g. for lattices) or not even distances but just categorizations of edges and thus follow a probability *mass* function. For these cases one may proceed as follows. Assume that pair-wise distances occur in an ordinal manner and there are n categories where

m_i is the number of pairs of distance r_i with $i = 1 \dots n$ and strict ordering $r_i < r_j$ if and only if $i < j$. Then the normalization condition is given by

$$\sum_{i=1}^{n} m_i C_\kappa r_i^{-\kappa} = \frac{N \langle k \rangle}{2}$$

and thus

$$C_\kappa = \frac{N \langle k \rangle}{2 \sum_{i=1}^{n} m_i r_i^{-\kappa}}.$$

As for the continuous-valued case above, there will be excess probability which has to be redistributed to categories of higher distance. Define j as the last category meeting the inequality

$$\sum_{i=1}^{j} m_i \left[C_\kappa r_i^{-\kappa} - 1 \right] > 0.$$

Then there is $\sum_{i=1}^{j} m_i \left[C_\kappa r_i^{-\kappa} - 1 \right]$ edges left that have to be redistributed to the $j+1$-st category. Hence, the final connection probability is

$$P_i = \begin{cases} 1 & i \leq j \\ C_\kappa r_{j+1}^{-\kappa} + \frac{1}{m_{j+1}} \sum_{i'=1}^{j} m_{i'} \left[C_\kappa r_{i'}^{-\kappa} - 1 \right] & i = j+1 \\ C_\kappa r_i^{-\kappa} & i > j+1. \end{cases} \tag{B.3}$$

This result is important because it shows how PLSW connection probabilities can be chosen to construct SSMH networks: Following the definitions given in Sec. 4.1, r can be a function of the hierarchical distance ℓ between two nodes and the m_i correspond to the number of possible pairs per hierarchical layer. In order to thoroughly map the two descriptions, a connection for r and ℓ has to be found, as well as a relation between the two structural control parameters κ and ξ. A demonstration on how this can be achieved will be given in Sec. 4.4.1. The result Eq. (B.3) is furthermore used for the derivation of the connection probability of a PLSW network on a one-dimensional lattice as follows.

B.3 Average Medium Approximations

B.3.1 Time Evolution of Walker Distribution on SSMH Model

Following Eq. (4.3), the probability of two nodes to be connected will be denoted as $\tilde{P}_\ell$. In the context of the AMA, this means that every node is connected to every other node but edges are weighted with probability $P_\ell = \tilde{P}_\ell / \langle k \rangle$. On this network, a random walk with a sink at an arbitrary node v is investigated. When a walker is positioned at node u, the probability to jump to node v at hierarchical distance $d(u, v) \equiv d$ is

$$P_d = \frac{1}{B-1} \left(\frac{1-\xi}{1-\xi^L} \right) \left(\frac{\xi}{B} \right)^{d-1},$$

as reasoned above. In this context, the following scenarios happen.

1. The walker jumps to node v. This occurs with probability

$$P_1 = P_d.$$

The new distance to node v will be $d' = 0$.

2. The walker jumps to a node in a lower layer $\ell < d$. The new distance will still be $d' = d$. In each layer $\ell < d$, there will be $B^{\ell-1}(B-1)$ possible target nodes for the walker starting at node u. The probability for this scenario to happen is

$$P_2 = \sum_{\ell=1}^{d-1} P_\ell B^{\ell-1}(B-1) = \frac{1-\xi^{d-1}}{1-\xi^L}.$$

3. The walker jumps to a node in $\ell = d$, but misses the branch that v is in and the branch that u is in (because this event is covered in scenario 2). The new distance is still $d' = d$. There are $B^{d-1}(B-2)$ nodes for this event. Consequently, it happens with probability

$$P_3 = P_d B^{d-1}(B-2).$$

4. The walker jumps to a node in $\ell = d$, hits the branch that v is in, but misses v. The new distance is $d' < d$, for each d' with probability

$$P_4 = P_d B^{d'-1}(B-1).$$

5. The walker jumps to a node in $\ell > d$. Consequently, the new distance is $d' > d$. The probability of this happening is, for each d'

$$P_5 = P_{d'} B^{d'-1}(B-1).$$

The transition matrix from a walker being at distance d from the target to being at new distance $0 \leq d' \leq L$ from the target is hence

$$W_{d'd} = \begin{cases} \frac{1}{(B-1)}\frac{1-\xi}{1-\xi^L}\left(\frac{\xi}{B}\right)^{d-1}, & d' = 0 \ \wedge \ d > 0 \\ \frac{B^{d'-1}}{B^{d-1}}\xi^{d-1}\frac{1-\xi}{1-\xi^L}, & d' < d \ \wedge \ d > 0 \\ \frac{1-\xi^{d-1}}{1-\xi^L} + \frac{1-\xi}{1-\xi^L}\xi^{d-1}\frac{B-2}{B-1}, & d' = d \ \wedge \ d > 0 \\ \frac{1-\xi}{1-\xi^L}\xi^{d'-1}, & d' > d \ \wedge \ d > 0 \\ 0, & d = 0 \ \wedge \ d = 0 \\ 1, & d' = 0 \ \wedge \ d = 0. \end{cases} \tag{B.4}$$

B.3.2 Global Mean First Passage Time

Suppose that a random walk begins with a single walker on every node $u \neq v$. Then the total ratio of walkers that are absorbed up to a time t is

$$\begin{aligned} p_0(t) &= \frac{1}{B^L - 1}\sum_{d_s=1}^{L} B^{d_s-1}(B-1)p_0^{(d_s)}(t) \\ &= \frac{1}{B^L - 1}\sum_{d_s=1}^{L} B^{d_s-1}(B-1)\left(\boldsymbol{W}^t\right)_{0,d_s}. \end{aligned} \tag{B.5}$$

Another way to write down the ratio of walkers absorbed into the sink node at time t is

$$p_0^{(d_s)}(t) = 1 - \sum_{d=1}^{L} p_d^{(d_s)}(t),$$

since the probability of the walker being at any distance $0 \leq d \leq L$ is equal to one and conserved at all times. At this point of the derivation a few new quantities need to be introduced. As described above, the probability of being at distance d' at time t is

$$p_{d'}(t) = \sum_{d=0}^{L} P_{d'd} p_d(t-1).$$

However, for $d' > 0$, the column $d = 0$ does not contribute anything to the sum (since it is filled with zeros). Hence, a new vector $\tilde{\boldsymbol{p}}^{(d_{\text{start}})}$ is defined which is the former probability vector with the 0-th element removed and a new transition matrix $\tilde{\boldsymbol{W}}$, as the full transition matrix Eq. (5.1) with the 0-th column and row removed. One therefore has

$$\tilde{p}_d^{(d_{\text{start}})}(t) = \left(\tilde{\boldsymbol{W}}^t\right)_{d,d_{\text{start}}},$$

such that

$$p_0^{(d_s)}(t) = 1 - \sum_{d=1}^{L} \left(\tilde{\boldsymbol{W}}^t\right)_{d,d_s}.$$

In order to consider the contribution of all starting nodes, these results are combined with Eq. (B.5) to find

$$p_0(t) = 1 - \frac{B-1}{B^L - 1} \sum_{d_s=1}^{L} B^{d_s - 1} \sum_{d=1}^{L} \left(\tilde{\boldsymbol{W}}^t\right)_{d,d_s}.$$

The GMFPT can be calculated as

$$\begin{aligned} \tau &= \sum_{t=0}^{\infty} t\left[p_0(t) - p_0(t-1)\right] = \sum_{t=0}^{\infty} t\left[\mathbf{1}^T \tilde{\boldsymbol{W}}^{t-1} \tilde{\boldsymbol{b}} - \mathbf{1}^T \tilde{\boldsymbol{W}}^t \tilde{\boldsymbol{b}}\right] \\ &= \mathbf{1}^T \left[\mathbb{1} - \tilde{\boldsymbol{W}}\right]^{-1} \tilde{\boldsymbol{b}}. \end{aligned} \tag{B.6}$$

Note that this quantity is equal to, both, the global mean FPT and the pair-averaged FPT (that is, the mean global mean FPT), because in this AMA all nodes are equal. This result is similar to the result for arbitrary networks, where instead of $(\mathbb{1} - \tilde{\boldsymbol{W}})^{-1}$ one makes use of the inverse of the reduced unnormalized graph Laplacian, as shown in Sec. 3.1.4. However, using the layer approach, the matrix size can be reduced from $B^L - 1$ to L, a great reduction in degrees of freedom. This makes it possible to obtain analytical expressions for the global mean FPT. As examples, the following results for $L < 4$ where found

solving Eq. (B.6):

$$
\begin{aligned}
\tau_{L=1} &= B-1\,,\\
\tau_{L=2} &= \frac{(B-1)B(\xi+1)\left(B^2\xi+B\xi+B-\xi\right)}{(B+1)\xi(B\xi+B-\xi)}\,,\\
\tau_{L=3} &= \Big[(B-1)B^2\left(\xi^2+\xi+1\right)\times\\
&\qquad\times\Big(B^4\xi^2(\xi+1)+B^3\xi(\xi+1)+\\
&\qquad\quad+B^2(\xi+1)^2-B\xi\left(2\xi^2+3\xi+2\right)+\xi^2(\xi+1)\Big)\Big]\\
&\qquad\Big/\Big[\left(B^2+B+1\right)\xi^2(B\xi+B-\xi)\times\\
&\qquad\quad\times\left(B\left(\xi^2+\xi+1\right)-\xi(\xi+1)\right)\Big]. \qquad \text{(B.7)}
\end{aligned}
$$

C
Additional Flockwork Evaluations

In Chap. 7, the temporal network model of "Flockworks" is introduced, however several detailed analyses are shortened to increase comprehensibility. These analysis are shown here. In Sec. C.1, the temporal group size evolution of the model is derived. Sec. shows more detailed analysis results of the state-space trajectory inference method. Finally, Sec. shows how to derive the evolution equations of an SIS-model with infinite infection rate on the Flockwork model.

C.1 Group Evolution

In Sec. 7.1.2, sethe temporal evolution of group sizes is discussed. The detailed considerations of how those group sizes change considering the events that can take place are, however, rather tedious and do not provide any particular insight. Hence, they have been banned to this appendix and are presented in the following.

In a Flockwork process defined by the event rate γ and the reconnection probability P as introduced in Sec. 7.1, the change in group sizes can be found for a single time step for both reconnection and disconnection events using P as a single control parameter because the time scale can be fixed by γ^{-1} for a single event.

First, the reconnection events are discussed, which each node undergoes with probability P in a single time step. The initial condition is that the network state can be expressed by the group size vector $\boldsymbol{n}$ with $0 \leq g \leq N+1$ where each entry n_g contains the total number of completely connected groups of size g. The group function $\kappa(u) = \mathrm{Nei}(u) + \{u\}$ is used.[1] These are the possible reconnection events.

[1] Note that the groups are of course time dependent, so it should be written as $\kappa(u,t)$, but in the following, the group size's time dependence is omitted in the notation.

1. A node u detaches from its group $\kappa(u)$ with group size $g = |\kappa(u)|$. It then reattaches to another group of size g, which is not the same group as its original one. The number of groups of size $g, g-1$ and $g+1$ change as

$$\Delta n_g = -2, \qquad \Delta n_{g-1} = 1, \qquad \Delta n_{g+1} = 1.$$

2. A node u detaches from its group $\kappa(u)$ with group size $g = |\kappa(u)|$. It then reattaches to the same group of size g. The number of

groups of size g change as

$$\Delta n_g = 0.$$

3. A node u detaches from its group $\kappa(u)$ with group size $g = |\kappa(u)|$. It then reattaches to a group of size $g-1$,. The number of groups of size g and $g-1$ change as

$$\Delta n_g = 0, \qquad \Delta n_{g-1} = 0.$$

4. A node u detaches from its group $\kappa(u)$ with group size $g = |\kappa(u)|$. It then reattaches to any group of size $i \neq g$ and $i \neq g-1$. The numbers of groups change as

$$\Delta n_g = -1, \qquad \Delta n_i = -1, \qquad \Delta n_{g-1} = 1, \qquad \Delta n_{i+1} = 1.$$

From these events, the following changes in the number of groups of size $1 \leq g \leq N$ have to be considered, where the superscript (r) denotes a reconnection event.

$$\begin{aligned}
\Delta n_g^{(\mathrm{r})} = \quad & P^{(\mathrm{r})}\Big[|\kappa(u)| = g,\ |\kappa(v)| = g,\ \kappa(u) \neq \kappa(v)\Big] \times (-2) \\
& + P^{(\mathrm{r})}\Big[|\kappa(u)| = g-1,\ |\kappa(v)| = g-1,\ \kappa(u) \neq \kappa(v)\Big] \times (+1) \\
& + P^{(\mathrm{r})}\Big[|\kappa(u)| = g+1,\ |\kappa(v)| = g+1,\ \kappa(u) \neq \kappa(v)\Big] \times (+1) \\
& + P^{(\mathrm{r})}\Big[|\kappa(u)| = g,\ |\kappa(v)| \neq g,\ |\kappa(v)| \neq g-1\Big] \times (-1) \\
& + P^{(\mathrm{r})}\Big[|\kappa(u)| = g+1,\ |\kappa(v)| \neq g,\ |\kappa(v)| \neq g+1\Big] \times (+1) \\
& + P^{(\mathrm{r})}\Big[|\kappa(u)| \neq g,\ |\kappa(u)| \neq g-1,\ |\kappa(v)| = g-1\Big] \times (+1) \\
& + P^{(\mathrm{r})}\Big[|\kappa(u)| \neq g+1,\ |\kappa(u)| \neq g,\ |\kappa(v)| = g\Big] \times (-1).
\end{aligned} \tag{C.1}$$

Note that here, the acting node is named u and node that u connects to is named v. The probabilities are proportional to the probabilities that from all N nodes, node u is picked *and* it reconnects *and* node v is picked from the remaining $N-1$ nodes afterwards. They are hence given as

$$
\begin{aligned}
&P^{(\mathrm{r})}\Big[|\kappa(u)| = g,\ |\kappa(v)| = g,\ \kappa(u) \neq \kappa(v)\Big] = \\
&\quad = P\frac{gn_g}{N} \times \frac{g(n_g - 1)}{N-1} \\
&P^{(\mathrm{r})}\Big[|\kappa(u)| = g-1,\ |\kappa(v)| = g-1,\ \kappa(u) \neq \kappa(v)\Big] = \\
&\quad = P\frac{(g-1)n_{g-1}}{N} \times \frac{(g-1)(n_{g-1} - 1)}{N-1} \\
&P^{(\mathrm{r})}\Big[|\kappa(u)| = g+1,\ |\kappa(v)| = g+1,\ \kappa(u) \neq \kappa(v)\Big] = \\
&\quad = P\frac{(g+1)n_{g+1}}{N} \times \frac{(g+1)(n_{g+1} - 1)}{N-1} \\
&P^{(\mathrm{r})}\Big[|\kappa(u)| = g,\ |\kappa(v)| \neq g,\ |\kappa(v)| \neq g-1\Big] = \\
&\quad = P\frac{gn_g}{N} \times \left(1 - \frac{gn_g - 1}{N-1} - \frac{(g-1)n_{g-1}}{N-1}\right) \\
&P^{(\mathrm{r})}\Big[|\kappa(u)| = g+1,\ |\kappa(v)| \neq g+1,\ |\kappa(v)| \neq g\Big] = \\
&\quad = P\frac{(g+1)n_{g+1}}{N} \times \left(1 - \frac{gn_g}{N-1} - \frac{(g+1)n_{g+1} - 1}{N-1}\right) \\
&P^{(\mathrm{r})}\Big[|\kappa(u)| \neq g-1,\ |\kappa(u)| \neq g,\ |\kappa(v)| = g-1\Big] = \\
&\quad = P\left(1 - \frac{(g-1)n_{g-1}}{N} - \frac{gn_g}{N}\right) \times \frac{(g-1)n_{g-1}}{N-1} \\
&P^{(\mathrm{r})}\Big[|\kappa(u)| \neq g+1,\ |\kappa(u)| \neq g,\ |\kappa(v)| = g\Big] = \\
&\quad = P\left(1 - \frac{(g+1)n_{g+1}}{N} - \frac{gn_g}{N}\right) \times \frac{gn_g}{N-1}.
\end{aligned}
$$

Now, the events at which nodes disconnect and subsequently stay alone have to be considered. These happen to every node with probability $1 - P$.

1. A node u from group $\kappa(u)$ with group size $g = |\kappa(u)| > 1$ detaches. The change in group sizes is

$$\Delta n_g = -1 \qquad \Delta n_1 = +1 \qquad \Delta n_{g-1} = +1.$$

2. A node u of group size $g = 1$ detaches. This does not change the group size distribution.

$$\Delta n_g = 0.$$

The new probabilities for a change in group size g are

$$
\begin{aligned}
P^{(\mathrm{d})}\Big[|\kappa(u)| = g\Big] &= (1-P)\frac{gn_g}{N} \\
P^{(\mathrm{d})}\Big[|\kappa(u)| > 1\Big] &= (1-P)\frac{N - n_1}{N} = \frac{1-P}{N}\sum_{g=2}^{N} gn_g.
\end{aligned}
$$

So the change in the number of groups of size g in a single time step are

$$
\begin{aligned}
\Delta n_{g>1}^{(\mathrm{d})} &= P^{(\mathrm{d})}\Big[|\kappa(u)| = g\Big] \times (-1) + P^{(\mathrm{d})}\Big[|\kappa(u)| = g+1\Big] \times (+1) \\
\Delta n_1^{(\mathrm{d})} &= P^{(\mathrm{d})}\Big[|\kappa(u)| > 1\Big] \times (+1) + P^{(\mathrm{d})}\Big[|\kappa(u)| = 2\Big] \times (+1).
\end{aligned}
$$

Combining Eqs. (C.1) and (C.2) one finds the evolution equations

$$N\Delta n_1 = -2Pn_1 + 2Pn_2\frac{N-2}{N-1} + (1-P)(N-n_1) + (1-P)2n_2\,, \tag{C.3a}$$

$$\begin{aligned} N\Delta n_{1<g\leq N} = \frac{P}{N-1}\Big[&-2gn_g(N-g) \\ &+ n_{g-1}\left(N(g-1)-(g-1)^2\right) + \\ &+ n_{g+1}\left(N(g+1)-(g+1)^2\right)\Big] - \\ &- (1-P)gn_g + (1-P)(g+1)n_{g+1}. \end{aligned}$$

Furthermore, no groups of size $g = 0$ and $g = N+1$ can exist, so $\Delta n_0 = \Delta n_{N+1} = 0$. Using the identity $N = \sum_{g=1}^{N} g n_g$, the term $N - n_1$ in Eq. (C.3a) can be expressed as $N - n_1 = \sum_{g=2}^{N} g n_g$ and hence the temporal evolution of the expected number of groups of size g can also be expressed using a transition matrix $\boldsymbol{W}$ such that

$$\boldsymbol{n}(t+1) = \boldsymbol{W}\boldsymbol{n}(t), \tag{C.4}$$

where

$$W_{gg'} = \begin{cases} 1-2P/N & g=1,\ g'=1 \\ 2\frac{P}{N}\frac{N-2}{N-1} + \frac{4}{N}(1-P) & g=1,\ g'=2 \\ \frac{g'}{N}(1-P) & g=1,\ g'>2 \\ P\frac{N(g-1)-(g-1)^2}{N(N-1)} & 1<g<N,\ g'=g-1 \\ 1-2P\frac{g(N-g)}{N(N-1)} - (1-P)\frac{g}{N} & 1<g<N,\ g'=g \\ (1-P)\frac{g+1}{N} + P\frac{N(g+1)-(g+1)^2}{N(N-1)} & 1<g<N,\ g'=g+1 \\ P/N & g=N,\ g'=N-1 \\ P & g=N,\ g'=N \\ 0 & \text{otherwise.} \end{cases} \tag{C.5}$$

The transitions of this process are discrete but the inter-event time follows an exponential distribution with the relevant time scale given by the inverse event rate γ^{-1}. Hence, the temporal evolution of $\boldsymbol{n}$ can be expressed by the continuous-time ordinary differential equation

$$\partial_t \boldsymbol{n} = \gamma\left(\boldsymbol{W} - \mathbb{1}\right)\boldsymbol{n}.$$

These are the temporal evolution equations presented in the main text Sec. 7.1.2. In the following, the equilibrium state of this process will be determined by considering that the overall change in a single time step vanishes (hence $\Delta\boldsymbol{n} = 0$) and recursively solving Eqs. (C.3). Consider Eq. (C.3a) with $\Delta n_1 = 0$ for the number of single nodes (group size $g = 1$). This gives the equilibrium state of the number of groups of size $g = 2$ as

$$n_2^* = \frac{N(P-1) + n_1(1+P)}{2P(N-2)/(N-1) + 2(1-P)}.$$

Now for every $2 \leq g \leq N-1$ the equations given by $\Delta n_g = 0$ can be solved to obtain an equation for the equilibrium value n^*_{g+1} which only depends on the equilibrium value n^*_1, recursively given as

$$\begin{aligned} n^*_{g+1} = \Big[& g^2 P(2n^*_g - n^*_{g-1}) + \\ & + g\left((N+2)n^*_{g-1}P - n^*_g(NP+N+P-1)\right) - \\ & - (N+1)n^*_{g-1}P \Big] \Big/ \left[(g+1)(gP-N+1)\right]. \end{aligned}$$

This process can be repeated until one arrives at $g = N-1$ and so obtains an equation for n^*_N that only depends on n^*_1, i.e. it is of the form

$$n^*_N = a n^*_1 + b$$

with a, b determined by the recursion relation. Subsequently, one uses Eq. (C.3a) for $g = N$ to obtain

$$n^*_N = \frac{P}{(1-P)N} n^*_{N-1},$$

an equation which also only depends on n^*_N and n^*_1 considering the previous recursive equations such that $n^*_{N-1} = c n^*_1 + d$. Now, both relations can be equated to find

$$n^*_1 = \frac{d-b}{a-c}$$

and hence all values of n^*_g by applying the recursive equations. Here, a computer algebra system was used to find the values of n^*_g in this manner, which, after applying the normalization condition $N = \sum_{g=1}^{N} g n_g$, are given as

$$n^*_1 = N(1-P) \tag{C.6a}$$

$$n^*_{1<g<N} = \frac{1}{g} \frac{N!}{(N-g)!} \frac{1-P}{\prod_{j=1}^{g-1}\left[1 - Pj/(N-1)\right]} \left(\frac{P}{N-1}\right)^{g-1} \tag{C.6b}$$

$$n^*_N = (N-2)! \frac{P}{\prod_{j=1}^{N-2}\left[1 - Pj/(N-1)\right]} \left(\frac{P}{N-1}\right)^{N-1}. \tag{C.6c}$$

This result was derived under the assumption that the initial network can be described as a collection of components which are alle fully connected. Below, it will be shown that this equilibrium distribution will be approached for *any* initial condition.

Consider two group count vectors, one counting the number n_g of groups of size $1 \leq g \leq N$ which are completely connected, and one counting the number $\bar{n}_g$ of groups of size $3 \leq g \leq N$ which are *not* fully connected. Hence, the state vector $(\boldsymbol{n}, \bar{\boldsymbol{n}})^T$ can completely contain any network (even though the explicit composition of the groups counted in the vector $\bar{\boldsymbol{n}}$ are ignored). Analyzing complete groups, the following events can happen.

1. A node disconnects from a fully connected group. It hence will still be a fully connected group containing one node less. Furthermore, the disconnect node will stay alone with probability $1 - P$, itself forming a complete network of size $g = 1$.

2. A node connects to a fully connected group. It hence will still be a fully connected group containing one node more.

It thus becomes clear that once a group is counted within the vector $\boldsymbol{n}$, it will stay in this category. The transition matrix describing the changes in counts of complete groups will be called $\boldsymbol{W}'$ in the following (compared to Eq. (C.4)). Now, consider groups which are not complete and thus are counted in the vector $\bar{\boldsymbol{n}}$. The following events may happen.

1. A node connects to a non-fully connected group. It hence will still not be fully connected but contain one group more.

2. A node disconnects from a non-fully connected group. A distinction between two cases has to be made.

   (a) The remaining group is now either fully connected and counted in the vector $\boldsymbol{n}$ or it is still non-fully connected and counted in the vector $\bar{\boldsymbol{n}}$.

   (b) Due to a node connecting multiple components of this group, the original group splits in multiple groups and each of those are now either counted in the vector $\bar{\boldsymbol{n}}$ or the vector $\boldsymbol{n}$.

Transition rates from groups in $\bar{\boldsymbol{n}}$ to fully connected groups in $\boldsymbol{n}$ will be contained in the matrix $\bar{\boldsymbol{W}}_1$ which has non-zero values in the upper triangle (because non-connected groups of size g can influence the count of fully connected groups of any size $g' < g$ but no fully connected groups of size $g' \geq g$). Transition rates from groups in $\bar{\boldsymbol{n}}$ to other groups in $\bar{\boldsymbol{n}}$ will be contained in the matrix $\bar{\boldsymbol{W}}_2$. This matrix has non-zero values in the upper triangle, the diagonal and the first diagonal shifted to the lower left, because non-fully connected groups of size g can influence the count of non-fully connected groups of size $g' < g$, non-fully connected groups of size $g' = g + 1$ (by reconnection) or $g' = g$. Then, the full time evolution is given as

$$\begin{pmatrix} \boldsymbol{n}(t+1) \\ \bar{\boldsymbol{n}}(t+1) \end{pmatrix} = \begin{pmatrix} \boldsymbol{W}' & \bar{\boldsymbol{W}}_1 \\ 0 & \bar{\boldsymbol{W}}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{n}(t) \\ \bar{\boldsymbol{n}}(t) \end{pmatrix}.$$

The shape of this equation implies that for any initial condition $(\boldsymbol{n}(0), \bar{\boldsymbol{n}}(0))^T$ the group count of $\bar{\boldsymbol{n}}$ will feed into the group count of $\boldsymbol{n}$ until $\bar{\boldsymbol{n}} = 0$. When this point is reached, the full temporal evolution is determined by Eq. (C.4) and hence the equilibrium Eq. (7.2) is reached eventually.

One may compute the expected number of components and the expected size of those components from the group-size distribution. The mean group size is

$$\langle g \rangle = \frac{\sum_{g=1}^{N} g n_g}{\sum_{g=1}^{N} n_g} = \frac{\sum_{g=1}^{N} g n_g}{\langle c \rangle} = \frac{N}{\langle c \rangle}.$$

with $\langle c \rangle$ being the mean number of components. Using Eq. (C.6) one may find[2] the exact result

[2] Here, this was done using a computer algebra system.

$$\begin{aligned}\frac{\langle c\rangle}{N} = 1 - P + \frac{1}{N}\underbrace{\frac{(-1)^N P\Gamma(N-1)\Gamma\left(\frac{-N+P+1}{P}\right)}{\Gamma\left(\frac{-N+P+1}{P}+N-2\right)}}_{A_1} + \frac{1-P}{P^2}\Bigg[\\ + \underbrace{\frac{\Gamma\left(\frac{-N+P+1}{P}\right)\left((-1)^N(N-1)P^2\Gamma(N-1)\right)}{\Gamma\left(\frac{PN-N+1}{P}\right)N}}_{A_2} + \\ + \underbrace{\frac{\Gamma\left(\frac{-N+P+1}{P}\right)}{\Gamma\left(\frac{-N+2P+1}{P}\right)}}_{A_3} \times (N-1-P) \times \\ \times \underbrace{\left(P - \frac{N-1}{N}\psi\left(\frac{1-N}{P}\right) + \frac{N-1}{N}\psi\left(\frac{N(P-1)+1}{P}\right)\right)}_{A_4}\Bigg],\end{aligned}$$

where $\Gamma(x)$ is the Gamma function and $\psi(x) = \Gamma'(x)/\Gamma(x)$ is the polygamma function. This result is incredibly unhelpful. However, it is possible to conjecture an asymptotic result for large numbers of nodes $N \gg 1$, which is done separately for each of the terms above.

For the first two terms, a numerical test shows that the conjectures $\lim_{N\to\infty} A_1/N = 0$ and $\lim_{N\to\infty} A_2 = 0$ seem appropriate. For A_3, one finds

$$A_3 \overset{N\gg 1}{\longrightarrow} -P/(N-1).$$

In the last term, one might use the fact that $\psi(x) \approx \log x$ for $x \gg 1$. Even though the arguments are negative here and hence this approximation should not be valid, one finds.

$$\begin{aligned} A_4 \quad &\approx \quad P + \frac{N-1}{N}\log\left(\frac{N(P-1)+1}{P}\right) - \\ &\qquad -\frac{N-1}{N}\log\left(\frac{1-N}{P}\right) \\ &= \quad P + \frac{N-1}{N}\log\left(1 - P\frac{N}{N-1}\right) \\ &\overset{N\to\infty}{\longrightarrow} \quad P + \log(1-P). \end{aligned}$$

Summing up, this yields the asymptotic formula

$$\begin{aligned}\lim_{N\to\infty}\frac{\langle c\rangle}{N} &= 1 - P + \Big(P + \log(1-P)\Big)\frac{1-P}{P^2} \\ &\qquad \lim_{N\to\infty}\left(-\frac{P}{N-1}\right)(N-1-P) \\ &= 1 - P - P\Big(P + \log(1-P)\Big)\frac{1-P}{P^2} \\ &= \frac{P-P^2}{P} - \frac{P-P^2+(1-P)\log(1-P)}{P} \\ &= -\frac{1-P}{P}\log(1-P),\end{aligned}$$

which also means that

$$\lim_{N\to\infty} \langle g \rangle = -\frac{P}{(1-P)\log(1-P)}.$$

A randomly chosen node is part of a group of size g with probability $P_g = gn_g/N$ and consequently has degree $k = g-1$. Hence, the mean degree is given as

$$\begin{aligned}\langle k \rangle &= \sum_{k=0}^{N-1} kP_k = \frac{1}{N}\sum_{k=0}^{N-1} k(k+1)n_{k+1} = \frac{1}{N}\sum_{g=1}^{N}(g-1)gn_g \\ &= \frac{1}{N}\sum_g g^2 n_g - 1 = \frac{1}{N}\left(\sum_g n_g\right)\frac{\sum_g g^2 n_g}{\sum_g n_g} - 1 = \frac{\langle c \rangle}{N}\left\langle g^2 \right\rangle - 1 \\ &= \frac{\langle g^2 \rangle}{\langle g \rangle} - 1. \end{aligned} \tag{C.7}$$

C.2 *Efficient Flockwork Equilibrium Configuration Sampling Algorithm*

In order to accurately simulate dynamic systems on a Flockwork model, e.g. the spread of diseases, it is necessary to initiate such a simulation in an already equilibrated system, however, depending on the initial state and the number of nodes, it can take a considerable number of events until the system is equilibrated. Instead, an algorithm to sample equilibrium configurations without having to simulate the Flockwork events is proposed in the following. Note that in contrast to the sections above, a distinction in notation will be made between the discrete observable number n_g of g-sized groups and its corresponding mean $\langle n_g \rangle$ as well as the theoretical mean $n_g^\star$ as given by Eqs. (7.2).

The algorithm will work by sampling discrete numbers n_g of groups which meet the normalization condition $\sum_{g=1}^{N} gn_g = N$ such that the average $\langle n_g \rangle$ of the samples will approach Eqs. (7.2). To this end it is first necessary to find the actual discrete distributions of n_g for varying reconnection probability P. Following the simulations shown in Fig. 7.3, the equilibrium distribution of n_g was measured over 10^4 independent simulations for a small number of nodes $N = 12$ and three values of reconnection probability $P \in \{0.05, 0.5, 0.99\}$. The results are shown in Fig. C.1. One notices that due to the normalization condition, the number of single nodes $n_1 = N-1$ cannot be realized (because then only a single node would be left, which, by definition, would increase the number of single nodes to $n_1 = N$). For small P, this pattern is repeated in the way that a triangle is highly unlikely, hence $n_1 = N-3$ has low probability. However, the probability of two pairs occuring is higher, and as such the probability of $n_1 = N-4$ is higher again. This behavior is difficult to replicate with single model distributions for each g since this would require a multivariate generalized Dirichlet distribution. In the following, the distributions will be approximated by independent discrete distribution.

1: $\ell \leftarrow N$ (ℓ is the number of nodes left to distribute)
2: $V \leftarrow$ random permutation of $\{u : (u \in \mathbb{N}) \wedge (1 \leq u \leq N)\}$
3: $n_g \leftarrow 0 \;\; \forall 1 \leq g \leq N$
4: $n_g^\star \equiv n_g^\star(N, P)$
5: **while** $\ell > 0$ **do**
6: **if** shuffle **then**
7: $M \leftarrow$ permutation of $(\{g : (g \in \mathbb{N}) \wedge (1 \leq g \leq \ell)\})$
8: **else**
9: $M \leftarrow$ decreasingly sorted $(\{g : (g \in \mathbb{N}) \wedge (1 \leq g \leq \ell)\})$
10: **end if**
11: **for** $g \in M$ **do**
12: **if** $n_g^*(N, P) > 0$ and $\ell \geq m$ **then**
13: draw $n_g^{\text{new}} \sim \mathcal{D}(N \text{ div } g, n_g^\star)$
14: **if** $n_g^{\text{new}} > n_g$ **then**
15: $\Delta n_g \leftarrow n_g^{\text{new}} - n_g$
16: **if** $g \times \Delta n_g > \ell$ **then** $\Delta n_g \leftarrow \ell \text{ div } g$
17: **end if**
18: **else if** $\ell = 1$ and $g = 1$ **then**
19: $\Delta n_g \leftarrow 1$
20: **else**
21: $\Delta n_g \leftarrow 0$
22: **end if**
23: **end if**
24: **for** $i \in 1, \Delta n_g$ **do**
25: Add g-sized group of node indices $V_{\ell-g+1} \leq u \leq V_\ell$
26: $n_g \leftarrow n_g + 1$
27: $\ell \leftarrow \ell - g$
28: **end for**
29: **end for**
30: **end while**

Algorithm 1: Sampling approximate equilibrium configurations for a Flockwork model. Note that here, $\mathcal{D}(N \text{ div } g, n_g^\star)$ is a placeholder for either of the two distributions $\mathcal{B}(N \text{ div } g, n_g^\star/[N \text{ div } g])$ or $\mathcal{P}(n_g^\star)$. The algorithm can be adjusted to iterate through group sizes in a randomly permuted or decreasingly sorted way with the boolean variable 'shuffle'.

First, consider a Poisson distribution $n_g \sim \mathcal{P}(n_g^\star)$. As one can observe in Fig. C.1, the Poisson distributions seem to approximate the actual distributions satisfyingly for most cases. However, in the extreme cases, i.e. $P = 0.05$ and $g = 1$, or $P = 0.99$ and $g = N$, the Poisson distribution will substantially differ from the real distribution, because the number of occurrences of g-sized groups is naturally upper bounded by $n_g^{\max} = N \,\text{div}\, g$ which the Poisson distribution does not satisfy. Instead, one might consider a Binomial distribution $n_g \sim \mathcal{B}(n_g^{\max}, n_g^\star/n_g^{\max})$, which seems to provide a better approximation of the real distributions in the extreme values of P, albeit ignoring the differences imposed by the normalization condition $\sum_{g=1}^N g n_g = N$ as discussed above. For the algorithm, both distributions will be tested.

The algorithm aims at drawing configurations n_g which satisfy $\sum_{g=1}^N g n_g = N$ as well as $\langle n_g \rangle = n_g^\star$. This will be pursued as follows. Consider that initially there are N nodes available for distribution to groups and $n_g = 0$ groups of size g. Going through a permuta-

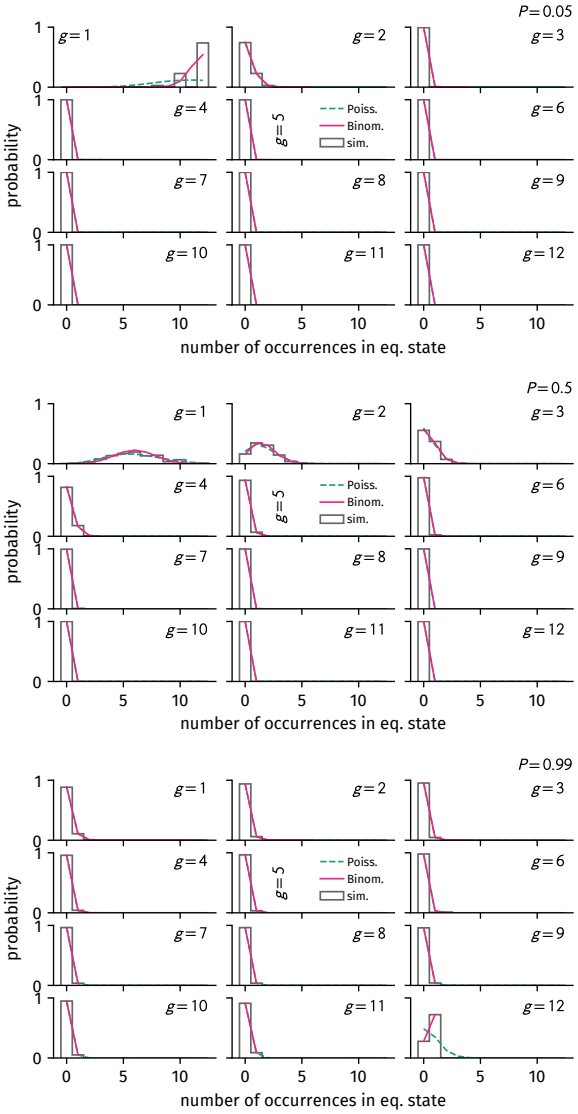


Figure C.1: The distribution of numbers n_g of g-sized groups in the equilibrium state for increasing reconnection probability P with **(top)** $P = 0.05$, **(middle)** $P = 0.5$, and **(bottom)** $P = 0.99$. Instead of analyzing the expected number of g-sized groups, here the full distributions are shown, as collected from the 10^4 final states of the simulations shown in Fig. 7.3. Additionally shown are Poisson and binomial distributions with the mean n_g^* from Eq. (7.2). For the binomial distributions, the number of trials to draw an g-sized group is given as N div g.

tion of all group sizes $1 \leq g \leq N$, for each g draw a number n_g^{new} from a distribution approximating the real distribution with mean n_g^* (i.e. either the binomial or the Poisson distribution) and add as many groups of size g which are still possible to add considering the number of nodes which are left. Update the new number n_g of g-sized groups accordingly. If, after all g have been visited, there are still nodes left, repeat this procedure. Now, only add Δn_g groups of size g where $\Delta n_g = \max\{0, n_g^{\text{new}} - n_g\}$. Repeat this process until all nodes are distributed to groups. This algorithm can be formalized as shown in Alg. 1.

A numerical analysis shows that the algorithm accurately constructs equilibrium configurations which are consistent with Eqs. (7.2) for a large number N of nodes and for $P < 1 - 1/N$ (cf. Fig. C.2). Both the binomial as well as the Poisson distribution seem to produce similar results. For a smaller number of nodes N, the decreasingly sorted way of group-size iteration yields more accurate configurations as otherwise smaller group sizes are increasingly selected.

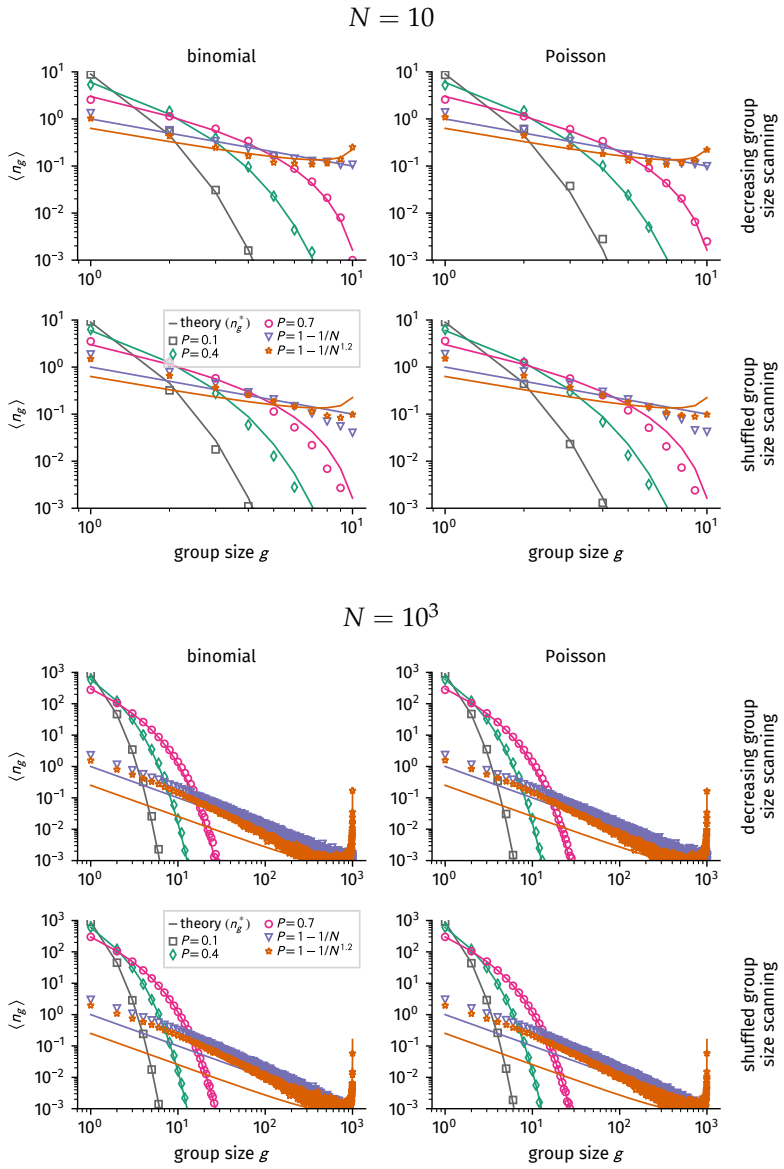


Figure C.2: Group size distribution as a mean over 10^4 independently constructed equilibrium configurations using Alg. 1 for varying values of P, **(top)** $N = 10$, **(bottom)** $N = 10^3$ nodes and different algorithmic details. For each panel these details are **(top row)** iterating over a decreasingly sorted list g of group sizes g, (bottom row) iterating over a shuffled list g of group sizes g, (left column) drawing the number n_g of g-sized groups from a binomial distribution for each g, and (right column) drawing the number n_g of g-sized groups from a Poisson distribution for each g. The curves obtained from the sampling algorithm are denoted by markers whereas solid lines represent Eqs. (7.2).

This changes with increasing number of nodes where the decreasingly sorted way of group-size iteration causes a biases to smaller group sizes. Hence, the shuffled way of group-size iteration is to be preferred. Furthermore, the sampled configurations become increasingly inaccurate with increasing reconnection probability P. For those values, a strong bias towards smaller group sizes exists while the shape of the distribution is nevertheless roughly reproduced.

When using the algorithm to sample equilibrium configurations it is therefore recommended to first scan all variants of the algorithm to see which of them best reproduce the true equilibrium group-size distributions for the chosen values of N and P. For large values of P one might improve the accuracy by sampling a configuration using Alg. 1 and then letting the sytem equilibrate to a true equilibrium configuration using the Flockwork procedure. The equilibration time for this procedure can be found by simulating a small number of

configurations and measuring the number of events until the group-size distributions correspond.

C.3 Link Decay Rate

For any parameter combination of $N = \{25, 50, 100\}$, and decreasing $0.999 > P > 0$, 100 Flockwork simulations were performed. Each of those was initiated using Alg. 1 with the options "binomial" and "shuffled group size iterations". Afterwards, the configuration was simulated for $T = 10^5(N\gamma)^{-1}$ to obtain a true equilibrium configuration. This equilibrium configuration was subsequently used as the initial configuration to simulate for a run time of $T = 2 \times 10^6(N\gamma)^{-1}$. From this simulation, the life-time was measured for active periods of each link which began neither at t_0 nor ended at $T + t_0$. Then $\langle \omega^- \rangle$ was obtained as the inverse of the mean of those life times and compared to Eq. (7.13). The results of this analysis are shown in Fig. C.3. One finds that Eq. (7.13) satisfyingly approximates the contact decay time with a maximum relative error of $< 2\%$ for $P = 0.999$ and $N = 100$. Furthermore, the relative error seems to decrease as a power-law as rel. err. $< A(1-P)^{-1}$ where A is a small constant. For larger numbers of N and higher values of P, it is however recommended to test the validity of Eq. (7.13) explicitly for the chosen values of N and P.

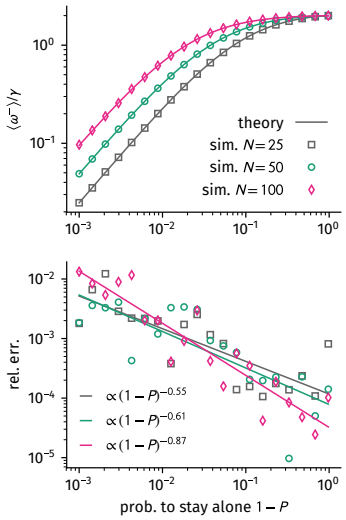


Figure C.3: Measurements vs. Eq. ((7.13)). Each data point is a mean over 10 independent simulations of runtime $T = 2 \times 10^6(N\gamma)^{-1}$.

C.4 Testing Rate-Inference with Synthetic Time-Varying Flockwork Data

The procedure described above is tested as follows. A system is defined in which a Flockwork model approaches several states in the (α, k_0)-plane by defining time-dependent rates $\alpha(t)$ and $\beta(t)$ which change their value every $\Delta t = 20$ as step functions. The unit of time will be fixed as $[t] = 1$ in the sense that if $\alpha(t) = 1$ then in one unit of time, there will be N active reconnection events on average. The rates will be loop over 4 consecutive state-changes and thus each loop is associated with a non-negative integer $\ell \geq 0$. The number of nodes was chosen as $N = 200$, the time-varying rates were set to

$$\alpha(t) = \begin{cases} 0.1 & 4\ell\Delta t \leq t < (4\ell+1)\Delta t \\ 10 & (4\ell+1)\Delta t \leq t < (4\ell+2)\Delta t \\ 10 & (4\ell+2)\Delta t \leq t < (4\ell+3)\Delta t \\ 0.1 & (4\ell+3)\Delta t \leq t < (4\ell+4)\Delta t, \end{cases} \tag{C.8}$$

and

$$\beta(t) = \begin{cases} 1 & 4\ell\Delta t \leq t < (4\ell+1)\Delta t \\ 100 & (4\ell+1)\Delta t \leq t < (4\ell+2)\Delta t \\ 1 & (4\ell+2)\Delta t \leq t < (4\ell+3)\Delta t \\ 0.1 & (4\ell+3)\Delta t \leq t < (4\ell+4)\Delta t, \end{cases} \tag{C.9}$$

such that, by using $P = \alpha/(\alpha + \beta)$ and $k_0 \approx P/(1 - P)$, the system is expected to move in the (α, k_0)-plane as $(0.1, 0.1) \to (10, 0.1) \to (10, 10) \to (0.1, 10) \to (0.1, 0.1)$ and so forth, as shown in Fig. C.4a. Essentially, the system should stay in each of those states for $\Delta t = 20$ and then transfer to the next state quickly, assuming that they equilibrate rapidly. Then, the parameters α, β, and k_0 should be correctly inferred using the procedure described above with $\Delta t_{\text{inference}} = \Delta t = 20$.

Instances of the Flockwork model with the defined input rates were simulated as follows. A Flockwork configuration was initiated using Alg. 1 with the options "Binomial distribution" and "shuffled group sizes" for $P = 0.1/(1 - 0.1)$ and $N = 200$. Then, the system was simulated using Gillespie's SSA with rates α and β as defined above until $t = 20$. Each time the rates switched, a new Flockwork simulation was initiated with the new rates and the last active configuration. The system was then simulated until $t = 40$ when the rates switched again. This procedure was repeated until $t = 320$, which means for four loops. Technically, this kind of simulation is "wrong" because when the rates switch, information about the time of the last event and the last rate has to be considered to evaluate the correct inter-event time (see App. D.2.3), this inter-event time when changing from one state to another, is, however, approximately upper bounded by $(N\gamma_{\min})^{-1} = (N\alpha_{\min} + N\beta_{\min})^{-1} \approx 1/(200 \times 0.2) = 1/40$, which is negligible compared to the time-interval length of the rate change $\Delta t = 20$.

Several instances of Flockwork simulations have been performed as described above, subsequently inferring the rates with varying inference parameter $\Delta t_{\text{inference}}$. The analysis of a single instance is displayed as an example in Fig. C.4 while a detailed analysis with more values of $\Delta t_{\text{inference}}$ is shown in App. F (see Figs. F.1 and F.2). As one may see in Fig. C.4b, the correct state-space trajectory is recovered using the input value $\Delta t = 20$ as the inference parameter. The rates $\alpha(t)$ and $\beta(t)$ are correctly inferred, too, as is shown in Fig. C.4d, alongside the curve of the instance's temporal variation of the configurational mean degree $\langle k \rangle (t)$ and the inferred expected mean degree $k_0(t)$, which is computed using the inferred rate functions and Eq. (7.16). Here, for each inference value $\Delta t_{\text{inference}}$, the expected mean degree $k_0(t)$ was evaluated for 10 equally distant points in each time-interval of length $\Delta t_{\text{inference}}$, hence for $\approx 3200/\Delta t_{\text{inference}}$ time points in total. In order to explore how the inference changes with different values of $\Delta t_{\text{inference}}$, a thorough scan with different values of $\Delta t_{\text{inference}}$was conducted, four of which are shown in Fig. C.4c and C.4e. Generally, values of $\Delta t_{\text{inference}} \leq \Delta t/2$ reproduce a fair picture of the state space trajectory but are subject to noise. The smaller the value of $\Delta t_{\text{inference}}$ is chosen, the more the rates will be fit directly to the equilibrium fluctuations, yielding overfitted results (see Fig. C.4e). When increased to values $\Delta t_{\text{inference}} > \Delta t$, the procedure yields heavily deviated space-state trajectories.

A thorough and quantitive method to find the optimal inference

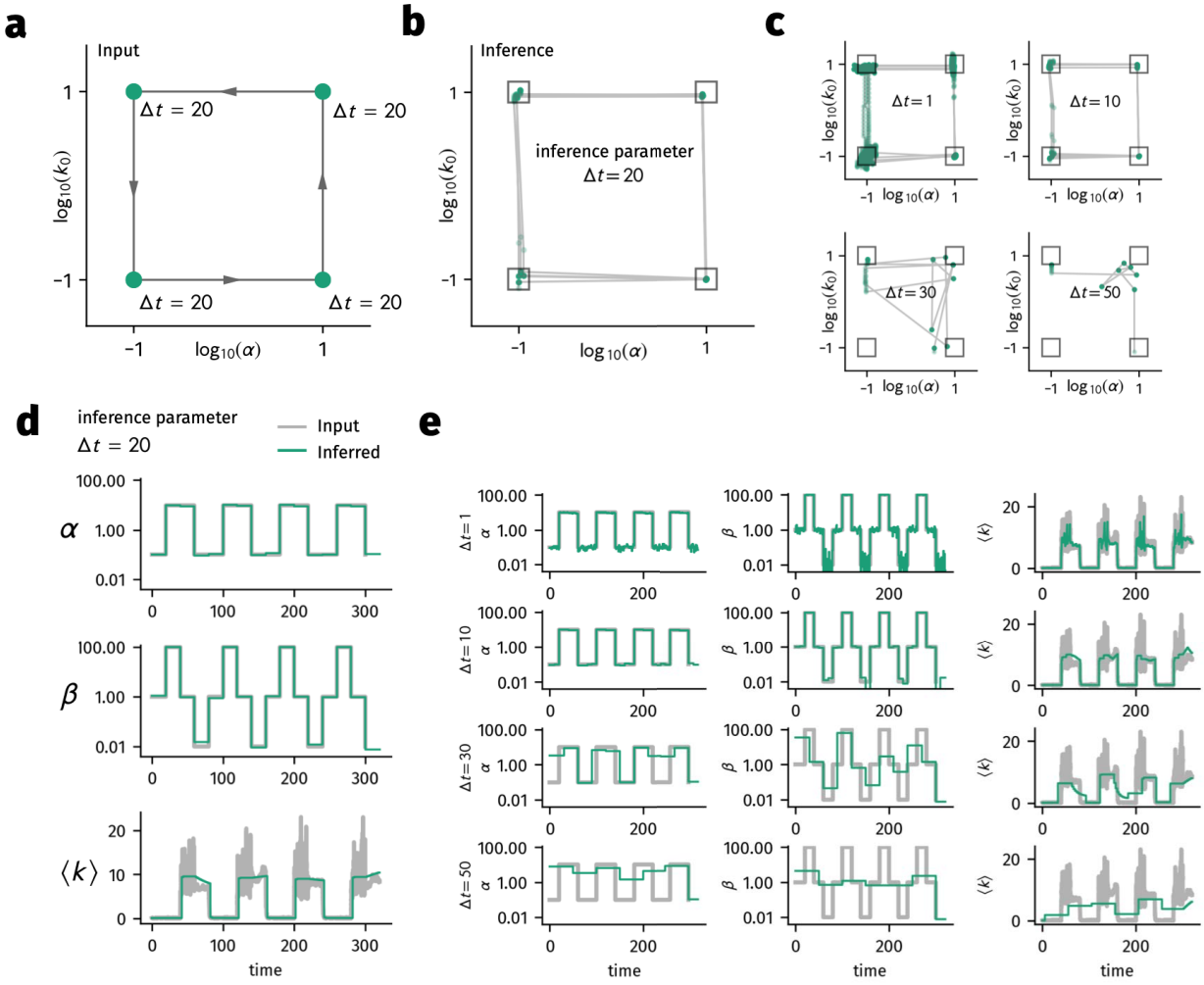


Figure C.4: Inferring time-varying rates and the state-space trajectory of a corresponding Flockwork simulation. **(a)** The input trajectory for the Flockwork simulation as based on the defined rate functions Eqs. (C.8) and (C.9). **(b)** The rate inference-procedure described in Sec. 7.2.1 yields the correct expected state-space trajectory. **(c)** Inference with time lengths smaller than the "true" Δt generally reproduce the "correct" state-space trajectory but are subject to noise. Inference using values of Δt greater than the "true" Δt will heavily deviate from the expected state-space trajectory. **(d)** The rates and expected mean degree are recovered correctly with only minor deviations when using the true interval length as the inference parameter. (e) Using $\Delta t_{\text{inference}} \leq \Delta t/2$ will yield rates fluctuating around the true functions, influenced by overfitting the fluctuations of the Flockwork's mean degree in its equilibrium state. Values of $\Delta t_{\text{inference}} > \Delta t$ will fail to accurately infer the rate functions and the expected mean degree.

parameter $\Delta t_{\text{inference}}$ without prior knowledge about the input functions will have to be the subject of future research. For real data, there is no reason for the time-varying rate functions to be step functions changing every Δt, so there is no *true* Δt. Rather, one may learn from this analysis that it may be of value to steadily increase the inference parameter and observe how the trajectories behave. One may then choose a Δt which reduces the noise from the trajectories obtained from small $\Delta t_{\text{inference}}$while striking features from the noisy trajectories are still present. Furthermore, it seems appropriate to compare the measured curve of $\langle k \rangle (t)$ to the inferred function $k_0(t)$ and choose a Δt which satisfyingly reproduces important features of the original curve while smoothing out heavy fluctuations. This method is used in the following to obtain the state-space trajectories of all data sets introduced in Sec. 2.3.3.

C.5 Group Size Evolution of Infected for Infinite Infectiousness

Here, a special case of an SIS infection process will be considered, where diseases are associated with an infinite infection rate $\eta \to \infty$ and a finite infection rate ρ. This situation implies that whenever an infected node connects to a susceptible node, it will infect every susceptible in a group before any other structural event may happen. Thus, any reconnection event between an infected and any group of susceptibles is likewise an infection event (as well as every reconnection event between a susceptible and group of infected, but less severe since only the reconnected susceptible node becomes infected). Hence, the probabilities for group size evolution of infected groups by re- and disconnection only within infected groups are analogous to the evolution before, with the probabilities

$$
\begin{aligned}
&P_I^{(\mathrm{r})}\Big[|\kappa(u)| = g,\ |\kappa(v)| = g,\ \kappa(u) \neq \kappa(v)\Big] \\
&\quad = P\frac{gI_g}{N} \times \frac{g(I_g - 1)}{N-1} \\
&P_I^{(\mathrm{r})}\Big[|\kappa(u)| = g-1,\ |\kappa(v)| = g-1,\ \kappa(u) \neq \kappa(v)\Big] \\
&\quad = P\frac{(g-1)I_{g-1}}{N} \times \frac{(g-1)(I_{g-1} - 1)}{N-1} \\
&P_I^{(\mathrm{r})}\Big[|\kappa(u)| = g+1,\ |\kappa(v)| = g+1,\ \kappa(u) \neq \kappa(v)\Big] \\
&\quad = P\frac{(g+1)I_{g+1}}{N} \times \frac{(g+1)(I_{g+1} - 1)}{N-1} \\
&P_I^{(\mathrm{r})}\Big[|\kappa(u)| = g,\ |\kappa(v)| \neq g,\ |\kappa(v)| \neq g-1\Big] \\
&\quad = P\frac{gI_g}{N} \times \left(\frac{I-1}{N-1} - \frac{gI_g - 1}{N-1} - \frac{(g-1)I_{g-1}}{N-1}\right) \\
&P_I^{(\mathrm{r})}\Big[|\kappa(u)| = g+1,\ |\kappa(v)| \neq g+1,\ |\kappa(v)| \neq g\Big] \\
&\quad = P\frac{(g+1)I_{g+1}}{N} \times \left(\frac{I-1}{N-1} - \frac{gI_g}{N-1} - \frac{(g+1)I_{g+1} - 1}{N-1}\right) \\
&P_I^{(\mathrm{r})}\Big[|\kappa(u)| \neq g-1,\ |\kappa(u)| \neq g,\ |\kappa(v)| = g-1\Big] \\
&\quad = P\left(\frac{I}{N} - \frac{(g-1)I_{g-1}}{N} - \frac{gI_g}{N}\right) \times \frac{(g-1)I_{g-1}}{N-1} \\
&P_I^{(\mathrm{r})}\Big[|\kappa(u)| \neq g+1,\ |\kappa(u)| \neq g,\ |\kappa(v)| = g\Big] \\
&\quad = P\left(\frac{I}{N} - \frac{(g+1)I_{g+1}}{N} - \frac{gI_g}{N}\right) \times \frac{gI_g}{N-1}. \\
&P_I^{(\mathrm{d})}\Big[|\kappa(u)| = g\Big] \\
&\quad = (1-P)\frac{gI_g}{N} \\
&P_I^{(\mathrm{d})}\Big[|\kappa(u)| > 1\Big] \\
&\quad = (1-P)\frac{1}{N}(g+1)I_{g+1}.
\end{aligned}
$$

Note that here, the total number of infected nodes in the network is given as

$$I = \sum_{g=1}^{N} g I_g.$$

Only considering reconnection events in between infected groups first (denoted by superscript (r)), the number of completely infected groups of size g would evolve as

$$\begin{aligned}
\gamma^{-1}\partial_t I_g^{(\mathrm{r})} = {} & P_I^{(\mathrm{r})}\Big[|\kappa(u)| = g,\ |\kappa(v)| = g,\ \kappa(u) \neq \kappa(v)\Big] \times (-2) \\
& + P_I^{(\mathrm{r})}\Big[|\kappa(u)| = g-1,\ |\kappa(v)| = g-1,\ \kappa(u) \neq \kappa(v)\Big] \times (+1) \\
& + P_I^{(\mathrm{r})}\Big[|\kappa(u)| = g+1,\ |\kappa(v)| = g+1,\ \kappa(u) \neq \kappa(v)\Big] \times (+1) \\
& + P_I^{(\mathrm{r})}\Big[|\kappa(u)| = g,\ |\kappa(v)| \neq g,\ |\kappa(v)| \neq g-1\Big] \times (-1) \\
& + P_I^{(\mathrm{r})}\Big[|\kappa(u)| = g+1,\ |\kappa(v)| \neq g,\ |\kappa(v)| \neq g+1\Big] \times (+1) \\
& + P_I^{(\mathrm{r})}\Big[|\kappa(u)| \neq g,\ |\kappa(u)| \neq g-1,\ |\kappa(v)| = g-1\Big] \times (+1) \\
& + P_I^{(\mathrm{r})}\Big[|\kappa(u)| \neq g+1,\ |\kappa(u)| \neq g,\ |\kappa(v)| = g\Big] \times (-1).
\end{aligned}$$

Second, nodes can detach, which yields the deconnection contributions

$$\begin{aligned}
\gamma^{-1}\partial_t I_{g>1}^{(\mathrm{d})} &= P^{(\mathrm{d})}\Big[|\kappa(u)| = g\Big] \times (-1) + P^{(\mathrm{d})}\Big[|\kappa(u)| = g+1\Big] \times (+1) \\
\gamma^{-1}\partial_t I_1^{(\mathrm{d})} &= P^{(\mathrm{d})}\Big[|\kappa(u)| > 1\Big] \times (+1) + P^{(\mathrm{d})}\Big[|\kappa(u)| = 2\Big] \times (+1).
\end{aligned}$$

Now, there are also infection-susceptible couplings that can happen.

1. The first chosen node is in group $\kappa(u) = g$ and infected. it chooses a susceptible group of size $\kappa(v) = n \neq g-1$. The change in infected group sizes is

$$\Delta I_g = -1 \qquad \Delta I_{n+1} = +1 \qquad \Delta I_{g-1} = +1$$

2. The first chosen node is in group $\kappa(u) = g$ and infected. it chooses a susceptible group of size $\kappa(v) = g-1$. The change in infected group sizes is

$$\Delta I_{g-1} = +1$$

3. The first chosen node is susceptible. It chooses an infected group of size $\kappa(v) = g$. The change in infected group sizes is

$$\Delta I_g = -1 \qquad \Delta I_{g+1} = +1$$

Thus, considering only reconnection events between infected and susceptible groups (denoted by superscript (i)), the number of com-

pletely infected groups of size g will evolve as

$$\begin{aligned}\gamma^{-1}\partial_t\left(I_g^{(\mathrm{i})}\right) &= P_I^{(\mathrm{i})}\Big[u\in\mathcal{I},|\kappa(u)|=g,v\notin\mathcal{I},\kappa(v)\neq g-1\Big]\times(-1)\\ &+ P_I^{(\mathrm{i})}\Big[u\in\mathcal{I},|\kappa(u)|=g+1,v\notin\mathcal{I}\Big]\times(+1)\\ &+ P_I^{(\mathrm{i})}\Big[u\in\mathcal{I},|\kappa(u)|\neq g,v\notin\mathcal{I},\kappa(v)=g-1\Big]\times(+1)\\ &+ P_I^{(\mathrm{i})}\Big[u\notin\mathcal{I},v\in\mathcal{I},\kappa(v)=g\Big]\times(-1)\\ &+ P_I^{(\mathrm{i})}\Big[u\notin\mathcal{I},v\in\mathcal{I},\kappa(v)=g-1\Big]\times(+1)\end{aligned}$$

with infection event probabilities

$$\begin{aligned}&P_I^{(\mathrm{i})}\Big[u\in\mathcal{I},|\kappa(u)|=g,v\notin\mathcal{I},\kappa(v)\neq g-1\Big]\\ &\quad= P\frac{gI_g}{N}\left(\frac{N-I}{N-1}-\frac{(g-1)(n_{g-1}-I_{g-1})}{N-1}\right)\\ &P_I^{(\mathrm{i})}\Big[u\in\mathcal{I},|\kappa(u)|=g+1,v\notin\mathcal{I}\Big]\\ &\quad= P\frac{(g+1)I_{g+1}}{N}\left(\frac{N-I}{N-1}\right)\\ &P_I^{(\mathrm{i})}\Big[u\in\mathcal{I},|\kappa(u)|\neq g,v\notin\mathcal{I},\kappa(v)=g-1\Big]\\ &\quad= P\left(\frac{I-gI_g}{N}\right)\frac{(g-1)(n_{g-1}-I_{g-1})}{N-1}\\ &P_I^{(\mathrm{i})}\Big[u\notin\mathcal{I},v\in\mathcal{I},\kappa(v)=g\Big]\\ &\quad= P\left(\frac{N-I}{N}\right)\frac{gI_g}{N-1}\\ &P_I^{(\mathrm{i})}\Big[u\notin\mathcal{I},v\in\mathcal{I},\kappa(v)=g-1\Big]\\ &\quad= P\left(\frac{N-I}{N}\right)\frac{(g-1)I_{g-1}}{N-1}.\end{aligned}$$

Here, n_g is the total number of groups of size g and hence the number of susceptible groups of size g is given as $n_g - I_g$. The group size evolution is happening in the same way as before. Additionally, every infected single node can recover with rate ρ and no node in an infected group $g > 1$ can recover, yielding

$$\begin{aligned}\partial_t I_1^{(\varrho)} &= -\rho\frac{I_1}{N}\\ \partial_t I_{g>1}^{(\varrho)} &= 0.\end{aligned}$$

Hence, the final evolution equations can be summed up as

$$\partial_t I_g = \partial_t I_g^{(\mathrm{r})} + \partial_t I_g^{(\mathrm{d})} + \partial_t I_g^{(\mathrm{i})} + \partial_t I_g^{(\varrho)},$$

while the total number n_g of groups of size g evolves as per Eqs. (7.1). This result highlights that the evolution of the number of infected is a combined effect of dis- and reconnection events within infected groups (d) and (r), meta-infection events (i) induced by reconnection events between susceptible and infected groups, and recovery events

(ρ) of single nodes. Eventually summing everything up yields the explicit equations of motion

$$\begin{aligned}
\partial_t I_1 = &- \underbrace{\frac{\rho I_1}{N}}_{\text{recovering}} + \underbrace{\frac{\gamma}{N}(1-P)(I-I_1+2I_2)}_{\text{nodes leaving infected groups}} + \\
&+ \underbrace{2\frac{\gamma P}{N(N-1)}\Big[I_2(I-2)-I_1(I-1)\Big]}_{\text{reconnections between infected}} + \\
&+ \underbrace{2\frac{\gamma P}{N(N-1)}(N-I)\,[I_2-I_1]}_{\text{reconnections between infected and susceptibles}} \\
\partial_t I_{g>1} = &\frac{\gamma P}{N(N-1)}\Big[(g-1)I_{g-1}(I-(g-1))+ \\
&\qquad + (g+1)I_{g+1}(I-(g+1)) - 2gI_g\,(I-g)\Big] \\
&+ \frac{\gamma}{N}(1-P)\Big[(g+1)I_{g+1} - gI_g\Big] \\
&+ \frac{\gamma P}{N(N-1)}\Bigg[I(g-1)\Big(n_{g-1}-I_{g-1}\Big)+ \\
&\qquad (N-I)\Big(-2gI_g + (g+1)I_{g+1} + (g-1)I_{g-1}\Big)\Bigg].
\end{aligned}$$

It is merely an approximation because the actual process is acting on the originally discrete states n_g and I_g. The equations above are ODEs acting on the mean of those discrete variables $\langle n_g \rangle$ and $\langle I_g \rangle$. In an exact formulation of the process, the ODEs would be of non-closed form including second-order correlations such as $\langle I \times I_g \rangle$ or $\langle I \times n_g \rangle$. Hence, the equations defined above are based on the assumption that these higher-order terms are approximately uncorrelated such that

$$\begin{aligned}
\langle I_g I_\ell \rangle &\approx \langle I_g \rangle \langle I_\ell \rangle \qquad \text{and} \\
\langle I_g n_\ell \rangle &\approx \langle I_g \rangle \langle n_\ell \rangle ,
\end{aligned}$$

which is assumed for any group sizes $1 \leq g \leq N$, $1 \leq \ell \leq N$. The equations of motion are tested for validity in Sec. E.2.

D
Gillespie's Stochastic Simulation Algorithm

Gillespie's stochastic simulation algorithm (SSA) is a general algorithm to simulate the evolution of Poisson processes related to discrete entities (e.g. particles, individuals, etc.). It was originally developed to simulate chemical reactions as birth-death processes in well-mixed systems. This Appendix introduces sequentially how and why the algorithm works for general processes with a single constant rate, multiple constant rates (for multiple events), single time-varying rates and multiple time-varying rates. It will close with an outline on how to perform exact simulations for locally constant time-varying rates.

D.1 Homogeneous Poisson Processes

This Section introduces Gillespie's SSA for processes happening with constant rates. For all of those processes consider the initial time to be $t_0 = 0$.

D.1.1 A Single Event

The simplest application for a stochastic simulation algorithm is one where only a single event can take place. This event happens with constant rate λ and is therefore a Poisson process. A Poisson process is based on several assumptions as stated in [140, pp. 59-60].

1. There is at maximum one event happening in an infinitesimal time interval of length dt.

2. The probability of one event happening in this time interval is λdt.

3. The process is Markovian in the sense that whether or not an event happens in this time interval does not depend on the events in any of the previous intervals.

Let $P_0(t)$ denote the probability that no event has taken place up to and including time t. According to assumptions 1 and 2, the proba-

bility that an event has happened in time dt is

$$P_1 = \lambda dt.$$

The probability that no event has taken place is consequently

$$P_0(dt) = 1 - P_1 = 1 - \lambda dt.$$

With Assumption 3 and due to the definition of P_0, the probability that no event has happened up to and including time $t + dt$ is hence the probability that no event has happened up to time t *and* no event happened within the subsequent time interval and thus

$$P_0(t + dt) = P_0(t) \times P_0(dt) = P_0(t) - P_0(t)\lambda dt$$

which translates to the ordinary differential equation

$$\frac{dP_0}{dt} = -\lambda P_0$$

with solution

$$P[\text{"no event happened"}, t] = P_0(t) = \exp(-\lambda t).$$

The complementary probability that at least one event has happened up to and including time t is consequently given as

$$\begin{aligned} P[\text{"at least one event happened"}, t] &= 1 - P_0(t) \\ &= 1 - \exp(-\lambda t) \\ &\equiv P(t) \end{aligned}$$

and the corresponding probability density for the time of the first event evaluates to

$$p(t) = \partial_t P(t) = \lambda \exp(-\lambda t).$$

Since this is just an exponential distribution, the algorithm is simply as follows. Draw a time t from the exponential distribution $t \sim \mathcal{E}(\lambda)$ and update the global time. Let the event take place and update the rate λ according to the new circumstances. Then start over with $t_0 = t$.

D.1.2 Multiple Events

Instead of one single event channel which can be triggered, in the following it is assumed that there are M such events, each taking place with constant rate $\lambda^{(i)}$, where $i \in [1, M]$.

Separate Evaluation

Since these processes are independent, one can draw a single potential leap time $t^{(i)}$ for each of the processes. Then, the process j where $t^{(j)} = \min\{t^{(i)}\}$ defines the new time $t \to t^{(j)}$ and the event which takes place is event j. Consequently, let the event take place, evaluate the new rates according to the new circumstances and start over with $t_0 = t$.

Joint Evaluation

Instead of evaluating each event channel on its own, one can evaluate the probability that none of the event channels have yet fired, which is

$$\begin{aligned} P[\text{"no event happened"}, t] &= P_0^{(1)}(t) \times \cdots \times P_0^{(M)}(t) \\ &= \prod_{i=1}^{M} \exp(-\lambda^{(i)} t) \\ &= \exp\left(-\sum_{i=1}^{M} \lambda^{(i)} t\right) = \exp\left(-\Lambda t\right). \end{aligned}$$

Consequently, the probability that at least one event has happened up to time t is given as

$$\begin{aligned} P[\text{"at least one event happened"}, t] &= 1 - P[\text{"no event happened"}, t] \\ &= 1 - \exp(-\Lambda t) \\ &\equiv P_\Lambda(t). \end{aligned}$$

The corresponding probability density for the time of the first event happening evaluates to

$$p_\Lambda(t) = \partial_t P_\Lambda(t) = \Lambda \exp(-\Lambda t).$$

Thus, the new time can simply be drawn from this exponential distribution as $t \sim \mathcal{E}(\Lambda)$. However, it is now necessary to decide which channel fired first. As argued above, the probability $\pi_j(t) \times p_\Lambda(t)dt$ that event j is the event which takes place *and* the time of the event is time t is equal to the probability that $t^{(j)} = \min\{t^{(i)}\}$ (the probability that it happens first). For $M = 2$ this is given as

$$\begin{aligned} \pi_1(t) p_\Lambda(t) dt &= P[t^{(1)} = \min\{t^{(i)}\} \text{ and } t^{(1)} = t] = P[t^{(1)} \text{ and } t^{(2)} \wedge t^{(1)} = t] \\ &= dt\ p^{(1)}(t) \int_t^\infty dt^{(2)}\ p^{(2)}(t^{(2)}) \\ &= dt\ p^{(1)}(t) \left[P^{(2)}(\infty) - P^{(2)}(t)\right] \\ &= dt\ \lambda^{(1)} \exp(-t\lambda^{(1)}) \exp(-t\lambda^{(2)}) \\ &= \frac{\lambda^{(1)}}{\lambda^{(1)} + \lambda^{(2)}} dt\ [\lambda^{(1)} + \lambda^{(2)}] \exp\left(-t[\lambda^{(1)} + \lambda^{(2)}]\right), \end{aligned}$$

such that at any time t, the probability that channel $j = 1$ fired first (given that it happened at time t) is

$$\pi_1(t) = \pi_1 = \frac{\lambda^{(1)}}{\lambda^{(1)} + \lambda^{(2)}}$$

and equivalently for the second event channel $\pi_2 = \lambda^{(2)}/[\lambda^{(1)} + \lambda^{(2)}]$. For $M = 3$ the probability that event channel $j = 1$ is the first one to fire given that the firing time is t is the probability that its

firing time $t^{(1)}$ fulfills $t^{(1)} < t^{(2)} < t^{(3)}$ or $t^{(1)} < t^{(3)} < t^{(2)}$ such that

$$\begin{aligned}
\pi_1(t)p_\Lambda(t)dt &= P[t^{(1)} < t^{(2)} \text{ and } t^{(1)} = t] \times P[t^{(2)} < t^{(3)}] + \\
&\quad + P[t^{(1)} < t^{(3)} \text{ and } t^{(1)} = t] \times P[t^{(3)} < t^{(2)}] \\
&= dt\ p^{(1)}(t) \Bigg[\int_t^\infty dt^{(2)}\ p^{(2)}(t^{(2)}) \int_{t^{(2)}}^\infty dt^{(3)}\ p^{(3)}(t^{(3)}) + \\
&\qquad + \int_t^\infty dt^{(3)}\ p^{(3)}(t^{(3)}) \int_{t^{(3)}}^\infty dt^{(2)}\ p^{(2)}(t^{(2)}) \Bigg] \\
&= dt\ p^{(1)}(t) \Bigg[\frac{\lambda^{(2)}}{\lambda^{(2)} + \lambda^{(3)}} \exp\left(-t[\lambda^{(2)} + \lambda^{(3)}]\right) + \\
&\qquad + \frac{\lambda^{(3)}}{\lambda^{(2)} + \lambda^{(3)}} \exp\left(-t[\lambda^{(2)} + \lambda^{(3)}]\right) \Bigg] \\
&= \frac{\lambda^{(1)}}{\lambda^{(1)} + \lambda^{(2)} + \lambda^{(3)}} \times \\
&\quad \times dt\ [\lambda^{(1)} + \lambda^{(2)} + \lambda^{(3)}] \exp\left(-t[\lambda^{(1)} + \lambda^{(2)} + \lambda^{(3)}]\right)
\end{aligned}$$

where similar steps have been used as during the evaluation of the case $M = 2$. By considering that evaluations for $M > 3$ will take similar forms one sees that this result generalizes for any M such that

$$\pi_j = \frac{\lambda^{(j)}}{\sum_{k=1}^M \lambda^{(k)}} = \frac{\lambda^{(j)}}{\Lambda}.$$

To sum up, once the event time t is decided, the corresponding event channel j is found by simply drawing a random integer from the interval $[1, M]$, each with probability π_j. Consequently, the event takes place, the rates are updated accordingly and the process starts over with $t_0 = t$.

D.2 Inhomogeneous Poisson Processes

This section introduces Gillespie SSAs for processes associated with time-varying rates. For all of those processes consider the initial time to be $t_0 = 0$.

D.2.1 A Single Event

As for the homogeneous case, consider a stochastic simulation algorithm where again only a single event can take place, whereas now, this event happens with a time-varying rate $\lambda(t)$ and is therefore associated with an inhomogeneous Poisson process. The assumptions made in Sec. D.1.1 have to be updated by replacing original assumption 2 with

2. The probability of one event happening in the time interval $[t, t + dt]$ is $\lambda(t)dt$.

Let $P_0|_t^{t+dt}$ denote the probability that no event has taken place within this time interval such that the probability that no event has taken place up to and including time t would be denoted as $P_0(t) = P_0\Big|_0^t$. According to assumptions 1 and 2, the probability that an event has happened within the time interval $[t, t+dt]$ is

$$P_1\Big|_t^{t+dt} = \lambda(t)dt,$$

such that the probability that no event has happened is

$$P_0\Big|_t^{t+dt} = 1 - \lambda(t)dt.$$

The probability $P_0(t+dt)$ that no event has taken place up to time $t+dt$ is, again (by assumption 3)

$$P_0(t+dt) = P_0(t) \times P_0\Big|_t^{t+dt} = P_0(t) - P_0(t)\lambda(t)dt$$

which evaluates to the differential equation

$$\frac{dP_0}{dt} = -\lambda(t)P_0$$

with general solution

$$P[\text{"no event happened"}, t] = P_0(t) = \exp\left[-\int_0^t d\tau\ \lambda(\tau)\right].$$

Consequently, the complementary probability that at least one event has happened is given as

$$\begin{aligned} P[\text{"at least one event happened"}, t] &= 1 - P_0(t) \\ &= 1 - \exp\left[-\int_0^t d\tau\ \lambda(\tau)\right] \\ &\equiv P(t) \end{aligned} \tag{D.1}$$

with corresponding probability density for the time of the first event happening

$$p(t) = \partial_t P(t) = \lambda(t) \exp\left[-\int_0^t d\tau\ \lambda(\tau)\right].$$

Considering the variable

$$\theta(t) = \int_0^t d\tau\ \lambda(\tau) \tag{D.2}$$

and comparing with Eq. (D.1) it becomes clear that θ is distributed according to an exponential distribution with unit mean $\theta \sim \mathcal{E}(1)$. Hence, to find the new time t, draw a random number θ and solve Eq. (D.2) for t (this can be done numerically). Subsequently, let the event take place, update the rate function $\lambda(t)$ accordingly and repeat the process with $t_0 = t$.

D.2.2 Multiple Events

Instead of one single event channel that can be triggered, in the following it is assumed that there's M such events, each taking place with time-varying rate $\lambda^{(i)}(t)$, where $i \in [1, M]$.

Separate Evaluation

Again, since these processes are independent, one can compute a single potential leap time $t^{(i)}$ for each of the processes by drawing M exponentially distributed random numbers $\theta^{(i)} \sim \mathcal{E}(1)$ and solving Eq. (D.2) for $t^{(i)}$. Then, the process j where $t^{(j)} = \min\{t^{(i)}\}$ defines the new time $t \to t^{(j)}$ and the event which takes place is event j. Consequently, let the event take place, evaluate the new rate functions and start over with $t = t_0$.

Joint Evaluation

Similarly as for the homogeneous event, instead of evaluating each event channel on its own, one can evaluate the probability that none of the event channels have yet fired, which is

$$\begin{aligned} P[\text{"no event happened"}, t] &= P_0^{(1)}(t) \times \cdots \times P_0^{(M)}(t) \\ &= \prod_{i=1}^{M} \exp\left(- \int_0^t d\tau \; \lambda^{(i)}(\tau) \right) \\ &= \exp\left(- \int_0^t d\tau \sum_{i=1}^{M} \lambda^{(i)}(\tau) \right) \\ &= \exp\left(- \int_0^t d\tau \Lambda(\tau) \right). \end{aligned}$$

Consequently, the probability that at least one event has happened up to time t is given as

$$\begin{aligned} P[\text{"at least one event happened"}, t] &= 1 - P[\text{"no event happened"}, t] \\ &= 1 - \exp\left(- \int_0^t d\tau \Lambda(\tau) \right) \\ &\equiv P_\Lambda(t). \end{aligned}$$

Furthermore, the corresponding probability density for the time of the first event happening is

$$p_\Lambda(t) = \partial_t P_\Lambda(t) = \Lambda(t) \exp\left(- \int_0^t d\tau \Lambda(\tau) \right).$$

Again, defining the variable

$$\Theta = \int_0^t d\tau \; \Lambda(\tau) \tag{D.3}$$

the new time can be found by drawing a random variable $\Theta \sim \mathcal{E}(1)$ and solving for t. Now, given the time of the event t, an event channel to be triggered has to be chosen. The probability $\pi_j(t) \times p_\Lambda dt$ that

event j is the event which takes place (given that the time t which was drawn from p_Λ) is equal to the probability that $t^{(j)} = \min\{t^{(i)}\}$ (the probability that the event of channel j happens first). For $M = 2$ this is given as

$$\begin{aligned}
\pi_1(t) \times p_\Lambda dt &= P[t^{(1)} = \min\{t^{(i)}\} \text{ and } t^{(1)} = t] = P[t^{(1)} < t^{(2)} \text{ and } t^{(1)} = t] \\
&= dt\ p^{(1)}(t) \int_t^\infty dt^{(2)}\ p^{(2)}(t^{(2)}) \\
&= dt\ p^{(1)}(t) \left[P^{(2)}(\infty) - P^{(2)}(t)\right] \\
&= dt\ \lambda^{(1)}(t) \exp\left[-\int_0^t d\tau\ \lambda^{(1)}(\tau)\right] \times \\
&\qquad \times \exp\left[-\int_0^t d\tau\ \lambda^{(2)}(\tau)\right] \\
&= dt\ \lambda^{(1)}(t) \exp\left[-\int_0^t d\tau\ \left[\lambda^{(1)}(\tau) + \lambda^{(2)}(\tau)\right]\right].
\end{aligned}$$

Given Eq. (D.3) one finds

$$d\Theta = \Lambda(t)dt$$

and thus

$$\pi_1(t(\Theta)) \exp(-\Theta) d\Theta = \frac{\lambda^{(1)}(t(\Theta))}{\Lambda(t(\Theta))} \exp(-\Theta) d\Theta.$$

Hence, given a random variable $\Theta \sim \mathcal{E}(1)$ and time $t(\Theta)$ as the (potentially numerical) solution of Eq. (D.3), the time-dependent probability that channel $j = 1$ fired at this time is

$$\pi_1(t) = \frac{\lambda^{(1)}(t)}{\Lambda(t)}.$$

This result is generalizable to $M > 2$. To see why this is the case, consider $M = 3$ where

$$\begin{aligned}
\pi_1(t) \times p_\Lambda dt &= P[t^{(1)} < t^{(2)} \text{ and } t^{(1)} = t] \times P[t^{(2)} < t^{(3)}] + \\
&\quad + P[t^{(1)} < t^{(3)} \text{ and } t^{(1)} = t] \times P[t^{(3)} < t^{(2)}] \\
&= dt\ p^{(1)}(t) \Bigg[\int_t^\infty dt^{(2)}\ p^{(2)}(t^{(2)}) \int_{t^{(2)}}^\infty dt^{(3)}\ p^{(3)}(t^{(3)}) + \\
&\qquad + \int_t^\infty dt^{(3)}\ p^{(3)}(t^{(3)}) \int_{t^{(3)}}^\infty dt^{(2)}\ p^{(2)}(t^{(2)}) \Bigg] \\
&= dt\ p^{(1)}(t) \Bigg[\int_t^\infty dt^{(2)}\ p^{(2)}(t^{(2)}) \exp\left[- \int_0^{t^{(2)}} dt\ \lambda^{(3)}(t) \right] + \\
&\qquad + \int_t^\infty dt^{(3)}\ p^{(3)}(t^{(3)}) \exp\left[- \int_0^{t^{(3)}} dt\ \lambda^{(2)}(t) \right] \Bigg] \\
&= dt\ p^{(1)}(t) \times \\
&\quad \times \Bigg[\int_t^\infty d\tau\ \lambda^{(2)}(\tau) \exp\left[- \int_0^\tau dt'\ \left(\lambda^{(3)}(t') + \lambda^{(2)}(t') \right) \right] + \\
&\quad + \int_t^\infty d\tau\ \lambda^{(3)}(\tau) \exp\left[- \int_0^\tau dt'\ \left(\lambda^{(3)}(t') + \lambda^{(2)}(t') \right) \right] \Bigg] \\
&= dt\ p^{(1)}(t) \exp\left(- \int_0^t d\tau\ \left(\lambda^{(3)}(\tau) + \lambda^{(2)}(\tau) \right) \right) \\
&= dt\ \lambda^{(1)}(t) \exp\left[- \int_0^t d\tau\ \lambda^{(1)}(\tau) \right] \times \\
&\qquad \times \exp\left[- \int_0^t d\tau\ \left(\lambda^{(3)}(\tau) + \lambda^{(2)}(\tau) \right) \right] \\
&= dt\ \left(\frac{\lambda^{(1)}(t)}{\Lambda(t)} \right) \Lambda(t) \exp\left(- \int_0^t d\tau\ \Lambda(\tau) \right).
\end{aligned}$$

With increasing M one can see that the equations will take forms similar to the ones above such that one can apply recursive solutions and end with the same general equation. Hence, with $\Lambda(t) = \sum_{i=1}^M \lambda_i(t)$ and Eq. (D.3) one finds, again,

$$\pi_j(t) = \frac{\lambda^{(j)}(t)}{\Lambda(t)}. \tag{D.4}$$

Consequently, after finding the time leap t, choose a random integer j from the interval $[1, M]$ with corresponding probability $\pi_j(t)$.

D.2.3 Application to Locally Constant Rates

A special case of an inhomogeneous Poisson process is one where event rates are time-varying but locally constant functions such that

$$\begin{aligned}
\lambda^{(i)}(t) &= \lambda_n^{(i)}, \qquad t \in [t_n, t_{n+1}), \text{ and} \\
\Lambda(t) &= \Lambda_n, \qquad t \in [t_n, t_{n+1}),
\end{aligned}$$

which is the case when the rates depend on changes of discrete variables such as contacts in time-varying networks or when using zero order approximations of continuous functions. Then, the integral Eq. (D.3) is exactly solvable. Considering that the process starts at time t_0, define $\Theta_0 = 0$. Draw a random number $\tilde{\Theta} \sim \mathcal{E}(1)$. Now, iteratively solve the integral as

$$\Theta_{n+1} = \Theta_n + \Lambda_n(t_{n+1} - t_n)$$

until $\tilde{\Theta} \leq \Theta_{n+1}$. The time t corresponding to $\tilde{\Theta}$ consequently lies within the time interval $(t_n, t_{n+1}]$. Find the exact time of the event by solving the linear solution of this part of the integral such that

$$t = t_n + \frac{\tilde{\Theta} - \Theta_n}{\Lambda_n}.$$

Subsequently, choose an event channel by drawing a random integer j from the interval $[1, M]$, each with probability $\pi_j = \lambda_n^{(j)} / \Lambda_n$, let the event take place, evaluate the new intervals $[t_n, t_{n+1})$, the corresponding rates λ_n, and start over with $t_0 = t$.

E
SIS-Simulations on Temporal Networks

E.1 Finite Infection Rates

Simulations of the SIS model on both the Flockwork model as well as the edge activity model were performed in the following way. The time scale was fixed using the active reconnection rate $\alpha = 1$ and the number of nodes was chosen as $N = 10^4$. The control parameters c_0, k_0 and R_0 were varied.

In order to choose k_0, the reconnection probability P was chosen as $P \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}$, then k_0 was calculated exactly from Eq. (7.6) for the Flockwork model. The corresponding parameter value for the edge activity model was consequently given as $p = k_0/(N-1)$. For c_0, the recovery rate ρ was given as $\rho \in \{0.01, 0.1, 0.5, 1, 2, 5, 10\}$ (in units of $\alpha = 1$). From k_0 and $\alpha = 1$, the Flockwork event rate $\gamma = \alpha/P$ was calculated such that simulations can be performed using the γ-P-formulation of the model. The corresponding parameter of link decay rate ω^- for the edge activity model was subsequently calculated using Eq. (7.13). Finally, the basic reproduction number was varied as $R_0 \in [10^{0.5}, 10^2]$ with 30 values evenly sized in logarithmic scale with base 10.

For each combination of those parameters, 40 simulations were perfomed for both the Flockwork model and the edge activity model to find the ratio $I^\star$.

Concerning the Flockwork model, for each of those simulations, one Flockwork equilibrium configuration was sampled to be used as an initial configuration at $t = 0$ using Alg. 1 with the options "Poisson distribution" and "shuffled group sizes". Then, 100 nodes were randomly chosen to be infected initially. Subsequently, a homogeneous joint Gillespie-simulation was performed with Flockwork event rate γ, infection rate per SI-contact η and recovery rate per infected ρ.

Considering the edge activity model, for each of those simulations of any control parameter combination, one equilibrium configuration was chosen as a sample from the Erdős–Rényi model with N nodes and connection probability $p = k_0/(N-1)$. Then, 100 nodes were randomly chosen to be infected initially. Subsequently, a homoge-

Table E.1: Chosen values of the equilibration time t_{eq} for SIS simulations on both the Flockwork model as well as the edge activity model.

recovery rate ρ $[\alpha]$	repr. number R_0	eq. time t_{eq} $[(N\alpha)^{-1}]$
< 0.1	≥ 20	100
	$10 \leq R_0 < 20$	170
	$6 \leq R_0 < 10$	280
	$4 \leq R_0 < 6$	400
	$2.6 \leq R_0 < 4$	590
	$1.89 \leq R_0 < 2.6$	1000
	$1.36 \leq R_0 < 1.89$	2000
	$R_0 < 1.36$	4000
≥ 0.1	≥ 6	40
	$4 \leq R_0 < 6$	100
	$1.5 \leq R_0 < 4$	140
	$1 \leq R_0 < 1.5$	180 (if $k_0 \geq 1$)
		500 (if $k_0 < 1$)
	$R0 < 1$	500

neous joint Gillespie-simulation was performed with global link decay rate ω^-, link creation rate $\omega^+ = \omega^- p/(1-p)$, infection rate per SI-contact η and recovery rate per infected ρ.

Each system was equilibrated until a time $t = t_{eq}$ which is given in Tab. E.1. Afterwards, the measurement phase of the simulation was performed until $t = t_{eq} + 200(N\alpha)^{-1}$, sampling the value of $I^\star(t)$ and $R_0(t) = \langle k \rangle (t)\eta/\rho$ after each $\Delta t = (N\alpha)^{-1}$. Consequently, $I^\star$was calculated as an average over those 201 recorded values, as well as the corresponding value R_0 to account for possibly large fluctuations in the network structure during simulation. However, if the simulation ended with $I^\star = 0$ at any time between $t = 0$ and $t = t_{eq} + 200(N\alpha)^{-1}$ then the average was not built over the values collected during the measurement. Instead, the simulation was counted as resulting in $I^\star = 0$. For the edge activity model, simulations for $k_0 = k_0(P = 0.95)$ and $c_0 = 100$ were omitted because of the extensive runtime of the simulations.

E.2 Infinite Infection Rate

In order to test the validity of Eq. (8.4), concurrent Flockwork/SIS-simulations have been performed using Gillespie's stochastic simulation algorithm. To this end, the number of nodes was chosen to be $N = 100$ and the contact-renewal number was varied as $c_0 \in \{1/8, 1/4, 1/2, 1, 2, 4, 8\}$. The structural control parameter was chosen as $P \in \{0.1, 0.3, 0.6, 0.9\}$, corresponding to $k_0 \in \{0.11, 0.43, 1.5, 9\}$. Then, the system was initiated with group size distribution $n_g(0) = (N, 0, \ldots, 0)$ and infected group size distribution $I_g(0) = (N, 0, \ldots, 0)$, meaning that at the beginning of the simulation the whole system was set to only consist of single disconnected nodes, all of which were infected. The timescale was fixed using $\alpha = 1$ per node and an infection rate was chosen as $\eta = 10^9\alpha$ (approaching the infinite limit). Then, for each of those parameter combinations, 100 independent simulations have been run until $T = 3/(N\alpha)$. Likewise,

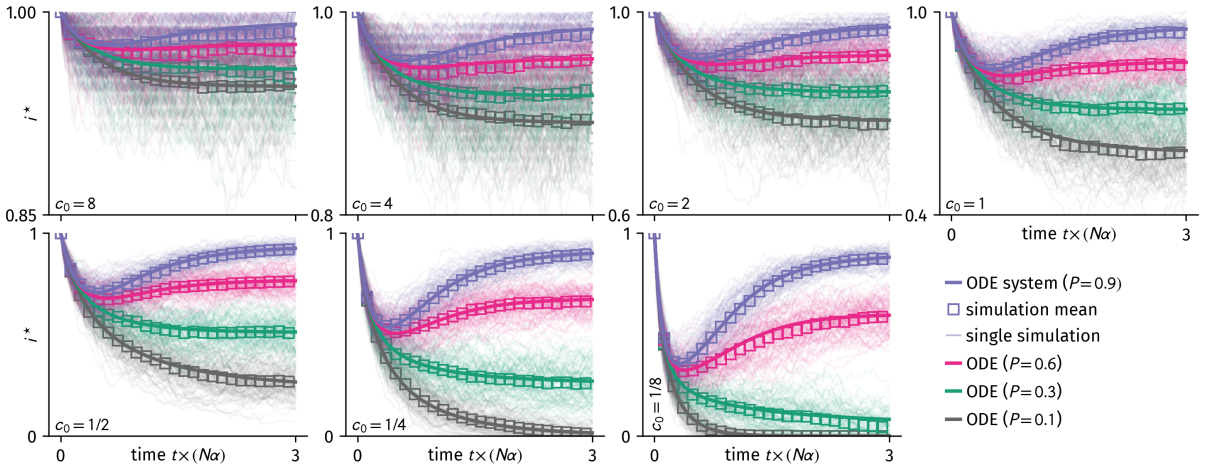


Figure E.1: Number $I^\star$ of infected for an infinitely infectious SIS process on the Flockwork model. Here, a system of $N = 100$ nodes was analyzed with constant active reconnection rate α, varying reconnection probability $P = \alpha/(\alpha + \beta)$ and varying contact-renewal number $c_0 = \alpha/\rho$. Shown are simulation results, means over simulation results, and the numerically found solution of Eqs. (7.1) and (8.4).

Eqs. (7.1) and Eqs. (8.4) have been integrated using a fourth-order Runge-Kutta integrator with step-size control. The results are presented in Fig. E.1. In general, the evolution functions obtained by integration are in agreement with the averages of the simulation results for all parameter values. For all larger mean degrees ($k_0 > 1$) as well as larger contact-renewal numbers ($c_0 > 1/4$) the processes seem to approach endemic equilibrium states which are $I^\star < N$. This will be further discussed in a separate analysis below. The functions show a behavior of initially falling number of infected over time with a subsequent phase of increasing number of infected for ensembles of higher mean degrees. This effect is caused by single infected nodes which initially recover but are then re-infected by new groups which are being formed. A further analysis showed that the same equilibrium states seem to be approached when the initial condition was set to be a complete graph of all infected nodes. Decreasing the total number of nodes does not seem to produce erroneous result either. However, Eqs. (8.4) yield wrongful results for a small number of nodes $N \lesssim 20$ when beginning with mixed initial conditions, e.g. a single lonely infected node. In this case correlations between the number of infected groups are becoming too high. In any case, the aim of this section is to find equilibrium numbers of infected. Using two different extreme initial conditions it was found that this is indeed possible by integrating Eqs. (8.4). Therefore, they can be used to find the endemic and disease-free states in the case of infinite infectiousness in Sec. 8.2.4.

E.3 Fitting Epidemic Response Curves

In Sec. 8.2.3, the hypothesis was made that the epidemic response curves obtained via Gillespie simulations follow a general law given by Eq. (8.2). In a second, more rigorous check for goodness of fit, the normalized residuals

$$r(R_0) = \frac{1}{\mathrm{Std}[i^\star](R_0)} \left[A - \frac{B}{(R_0 - D)^C} - i^\star_{\mathrm{sim}}(R_0) \right] \qquad \text{(E.1)}$$

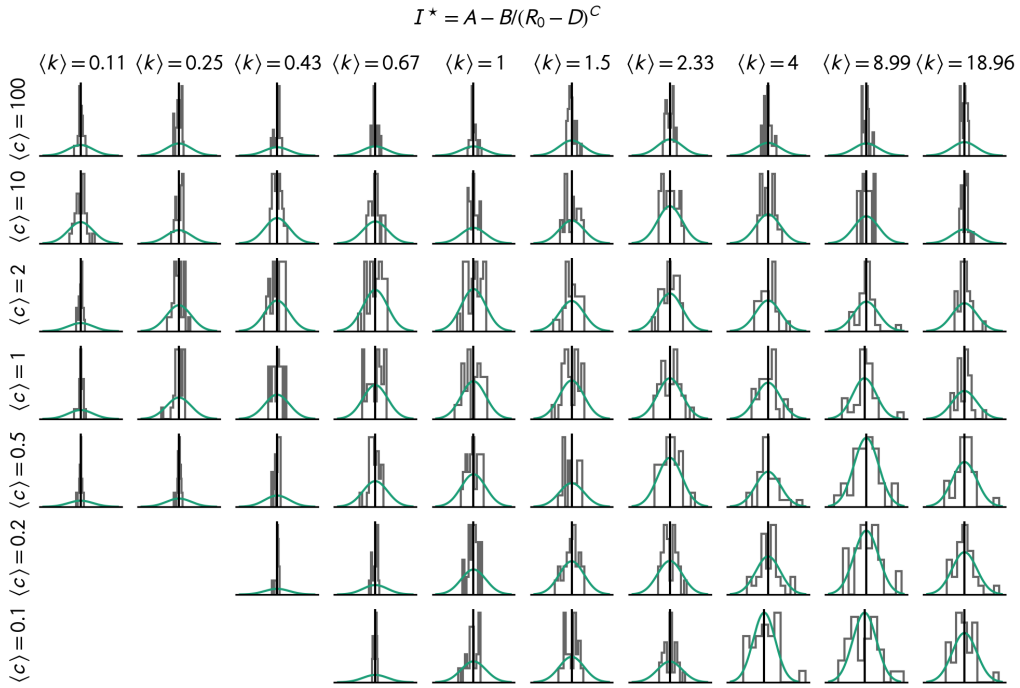


Figure E.2: Normalized residuals of the fits to the SIS simulation data on the Flockwork model in Sec. 8.2.3. Solid green lines represent the normal distribution $\mathcal{N}(0,1)$ and grey boxes are the histogram of the residuals as per Eq. (E.1). Note that here, $\langle c\rangle \equiv c_0$ and $\langle k\rangle \equiv k_0$.

were analyzed. For a successful fit, the residuals should follow a Gaussian distribution of zero mean and unit standard deviation[141]. These comparisons are shown in Fig. E.2 for the Flockwork model and in Fig. E.3 for the edge activity model. The law Eq. (8.2) often overfits the simulation data often but does not significantly underfit. More simulation results would be necessary to make proper statistically sound statements about the validity of Eq. (8.2). However, the purpose of the law is to find the epidemic threshold which appears to be estimated sufficiently well by Eq. (8.3) when compared to the simulation results. In order to test the influence of overfitting, each law of a subset of the parameters A, B, C, and D was tested as a model for the simulation data (with corresponding fixed mean-field values) and the best model was chosen to obtain the epidemic threshold.[1] Since these evaluations did not yield results significantly different from using the full model, the full four-parameter model will be used.

[1] Here, "best model" refers to the model with minimal Kolmogorov-Smirnov-distance between residual distribution and normal distribution as suggested in [141].

E.4 Simulations on Real-World Data

E.4.1 Methods

After careful evaluations of single infection trajectories, an equilibration time of $t_{\text{eq}} = 700\,\text{d}$ was chosen for all Gillespie simulations and Markov-chain integration. Longer equilibration times were tested, especially for low basic reproduction numbers, but were found to not

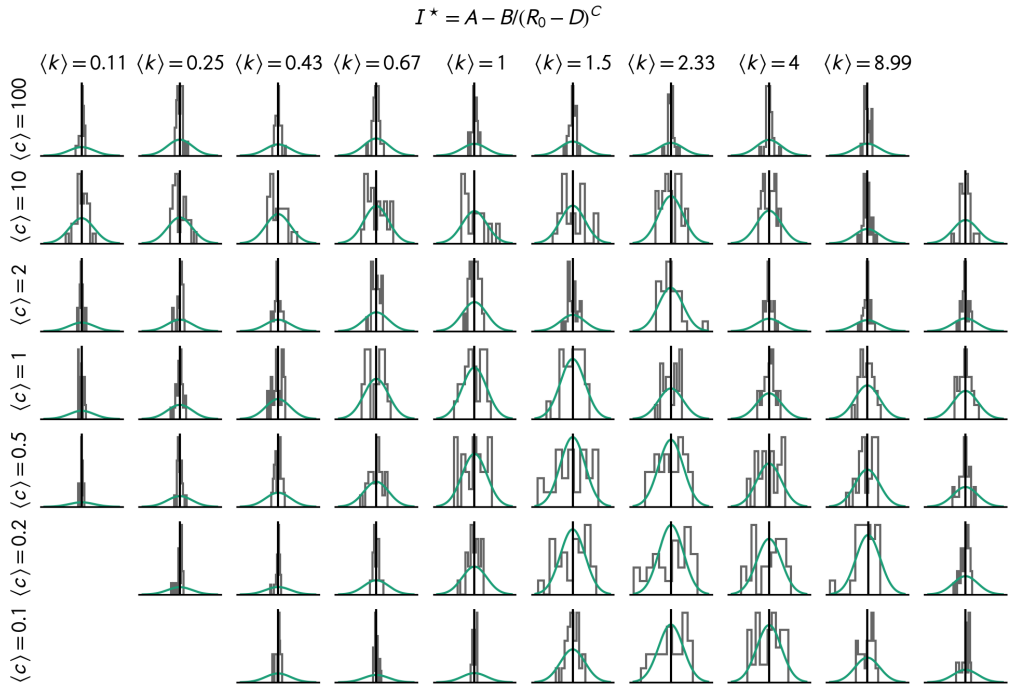


Figure E.3: Normalizedresiduals of the fits to the SIS simulation data on the edge activity model in Sec. 8.2.3. Solid green lines represent the normal distribution $\mathcal{N}(0,1)$ and grey boxes are the histogram of the residuals as per Eq. (E.1). Note that here, $\langle c\rangle \equiv c_0$ and $\langle k\rangle \equiv k_0$.

change the response curves. If not indicated otherwise, the subsequent measurement period was chosen as $t_{\text{meas}} = 70\,\text{d}$ with samples of the total number of inftected nodes taken every hour. The basic reproduction number was varied between $10^{-0.5} \leq R_0 \leq 10^{1.5}$ with 49 logarithmically spaced steps. While $N_{\text{meas}} = 100$ independent simulations where performed with the Gillespie method for each parameter combination, a single integration of the Markov-chain approximation equations Eq. (3.26) was performed with $\Delta t_{\text{max}}^{\text{DTU}} = 5\,\text{min}$, $\Delta t_{\text{max}}^{\text{HT09}} = \Delta t_{\text{max}}^{\text{HS13}} = 20\,\text{s}$. Sample curves with smaller values of Δt_{max} yielded the same results and hence the computationally cheaper value was chosen. Note that instead of discrete infection states, The Markov-integration yields the probability $\langle i_u\rangle(t)$ of a node u to be infected such that $I(t) = \sum_{u=1}^{N}\langle i_u\rangle(t)$ is the average total number of infected. The integration was stopped if $I(t) \leq 10^{-6}$ at which $I^\star = 0$ was assumed, or if $t = t_{\text{eq}} + t_{\text{meas}}$ after which $i^\star$ was evaluated using Eq. (8.6).

To simulate on static averaged networks, the Gillespie simulation algorithm introduced in Sec. 3.2.3 is adapted to work on symmetric, weighted networks with entries $w_{uv} = w_{vu}$ and $w_{uv} \in [0,1]$ where $w_{uv} = 0$ means that nodes u and v never have contact and $w_{uv} = 1$ means that nodes u and v are *always* connected. The total event rate of the Gillespie simulation is then given as

$$\Lambda(t_0) = \rho \sum_{u=1}^{N} i_u(t_0) + \eta \sum_{u=1}^{N}\sum_{v=1}^{N} i_u(t_0)[1 - i_v(t_0)]w_{uv}$$

where here, $i_u = 1$ if node u is infected and $i_u = 0$ if not. In case

a recovery event is evaluated to happen, an infected node will be chosen randomly from the set of infected nodes and recovers. If an infection event is chosen, each susceptible node is associated with a force of infection acting on it, given as $f_v = (1 - i_v) \sum_{u=1}^{N} i_u w_{uv}$. Consequently, the node v to be infected is chosen randomly with probability $\pi^{(v)} = f_v / \sum_{u=1}^{N} f_u$ from the set of all susceptible nodes.

E.4.2 Circadian Influence Analysis

In order to compare the influence of circadian activity on epidemic response curves, additional Gillespie simulations with $N_{\text{meas}} = 1000$ were performed and the endemic state was evaluated when at least $n = 10$ of those simulations did not end in the absorbing state. As substrates, all three real-world data sets DTU, HT09, and HS13 were used as well as their corresponding Flockwork surrogates. The results for the real-world data sets are shown in Fig. E.4 and Fig. E.5, while the results for the FW data sets are presented in Fig. E.6 and Fig. E.7.

E.5 SIS-Model on a Pair

In [45], the exact temporal evolution of an SIS process on a single pair $\bullet - \bullet$ was derived. Their results show that the total number of infected approaches zero in equilibrium. They further evaluated a lower bound of an epidemic threshold, which is puzzling because there is no epidemic threshold in the classical sense: on a finite network, the absorbing state is always reached [142]. In order to illustrate problems with the individual-based Markov-chain approximation, an adapted SIS-model is evaluated in the following.

E.5.1 The ε-SIS-Model and the MCA

The ε-SIS-model is an adaption of the SIS-model in which a third reaction equation is added

$$S \xrightarrow{\varepsilon} I,$$

which describes events where a susceptible node is spontaneously infected by itself [142], modeling a constant influx of the disease. As will be shown in the next section, the constant influx allows for an endemic stationary state to establish.

The individual-based Markov-chain approximation for a single pair of nodes reads

$$\begin{aligned} \partial_t \langle i_1 \rangle &= \varepsilon(1 - \langle i_1 \rangle) - \rho \langle i_1 \rangle + \eta(1 - \langle i_1 \rangle) \langle i_1 \rangle \\ \partial_t \langle i_2 \rangle &= \varepsilon(1 - \langle i_2 \rangle) - \rho \langle i_2 \rangle + \eta(1 - \langle i_2 \rangle) \langle i_2 \rangle . \end{aligned} \tag{E.2}$$

Hence, the Jacobian matrix in the stable state $\langle i_1 \rangle = \langle i_2 \rangle = 0$ has maximum eigenvalue $j_{\max} = \eta - \varepsilon - \rho$. With the definitions introduced above and noticing that the epidemic threshold is located at

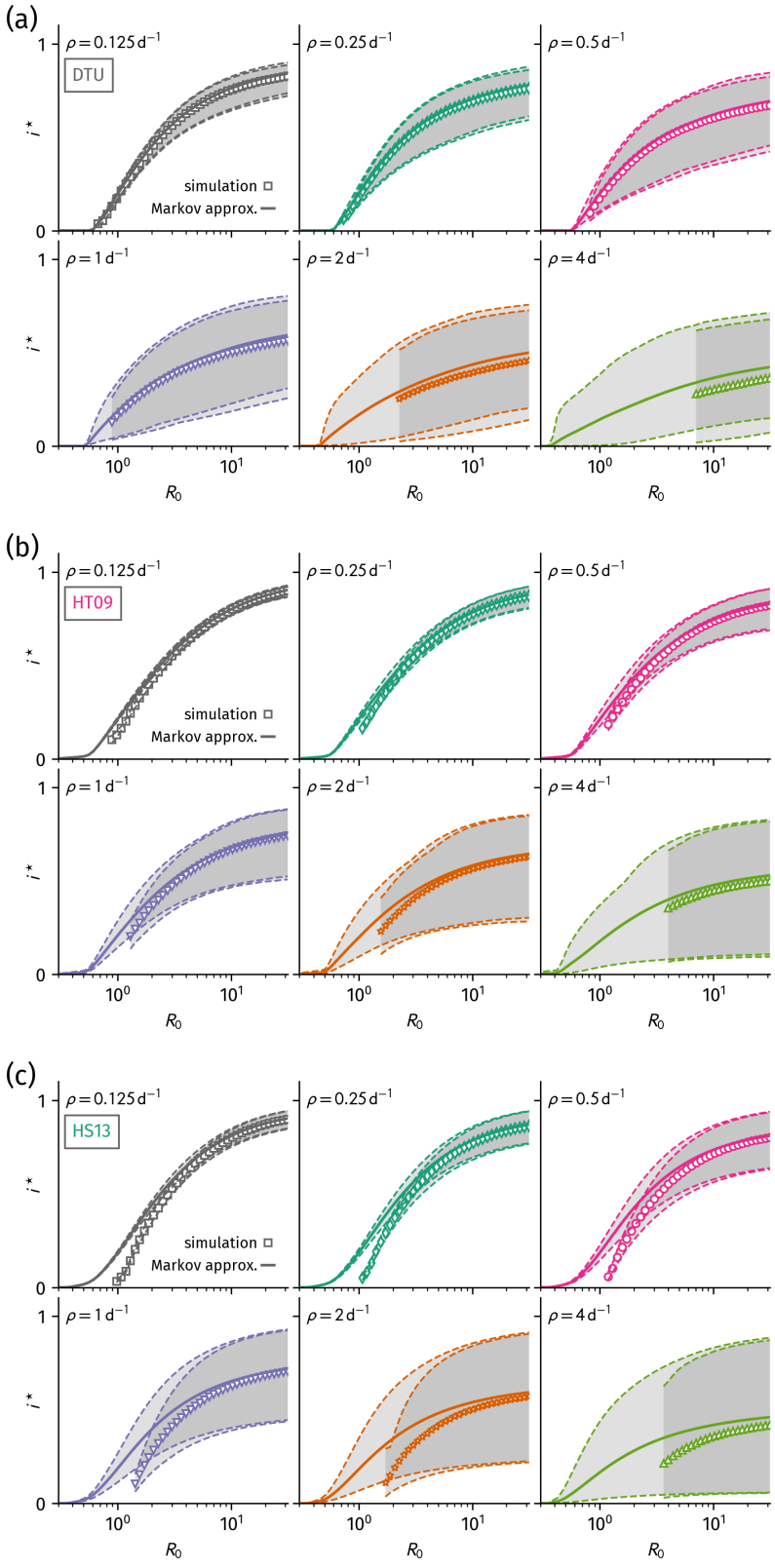


Figure E.4: The influence of circadian activity on epidemic response curves and epidemic thresholds, as a comparison between Gillespie simulations and the individual-based Markov-chain approximation. Since the networks investigated here become sparse and slowly varying at night (or during the weekend), the endemic state potentially decreases strongly during those times. Solid lines and markers represent the average endemic state Eq. (8.6) and grey bands mark the extent of their variation in the . When the minimal number of infected $i^\star_{\min}$ (lower dashed lines) approaches $i^\star_{\min} = 0$, the disease becomes unsustainable and dies out. The average endemic state is potentially much larger than the minimal endemic state and therefore, sudden jumps in the order parameter $i^\star$ occur.

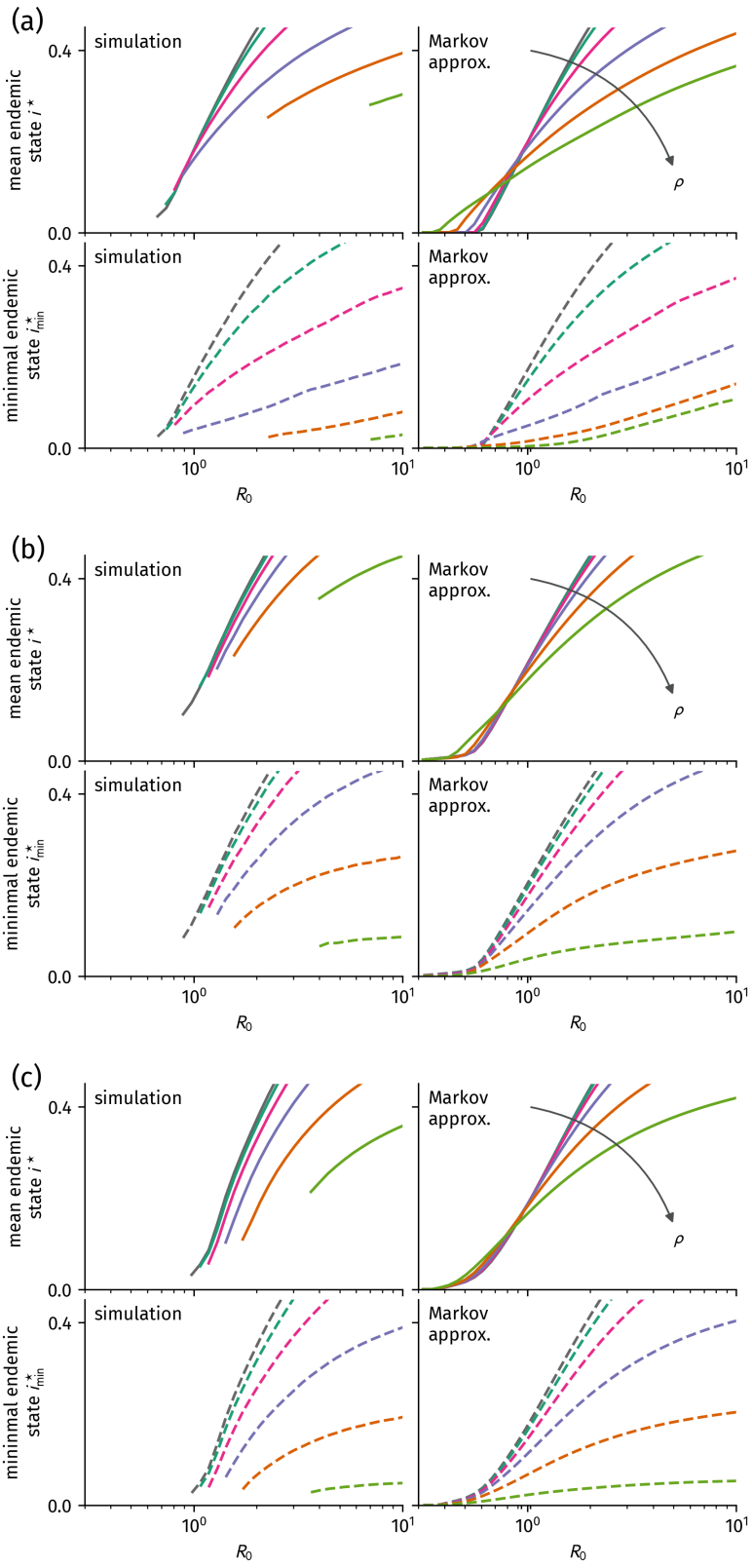


Figure E.5: Enlarged illustration of the influence of circadian activity on epidemic thresholds, as a comparison between Gillespie simulations and the individual-based Markov-chain approximation. The average minimal number of infected during an observation period is a more reliable order parameter to obtain the epidemic threshold. The Markov-chain approximation method gives increasingly differing predictions with increasing recovery rate ρ. (a) DTU, (b) HT09, (c) HS13.

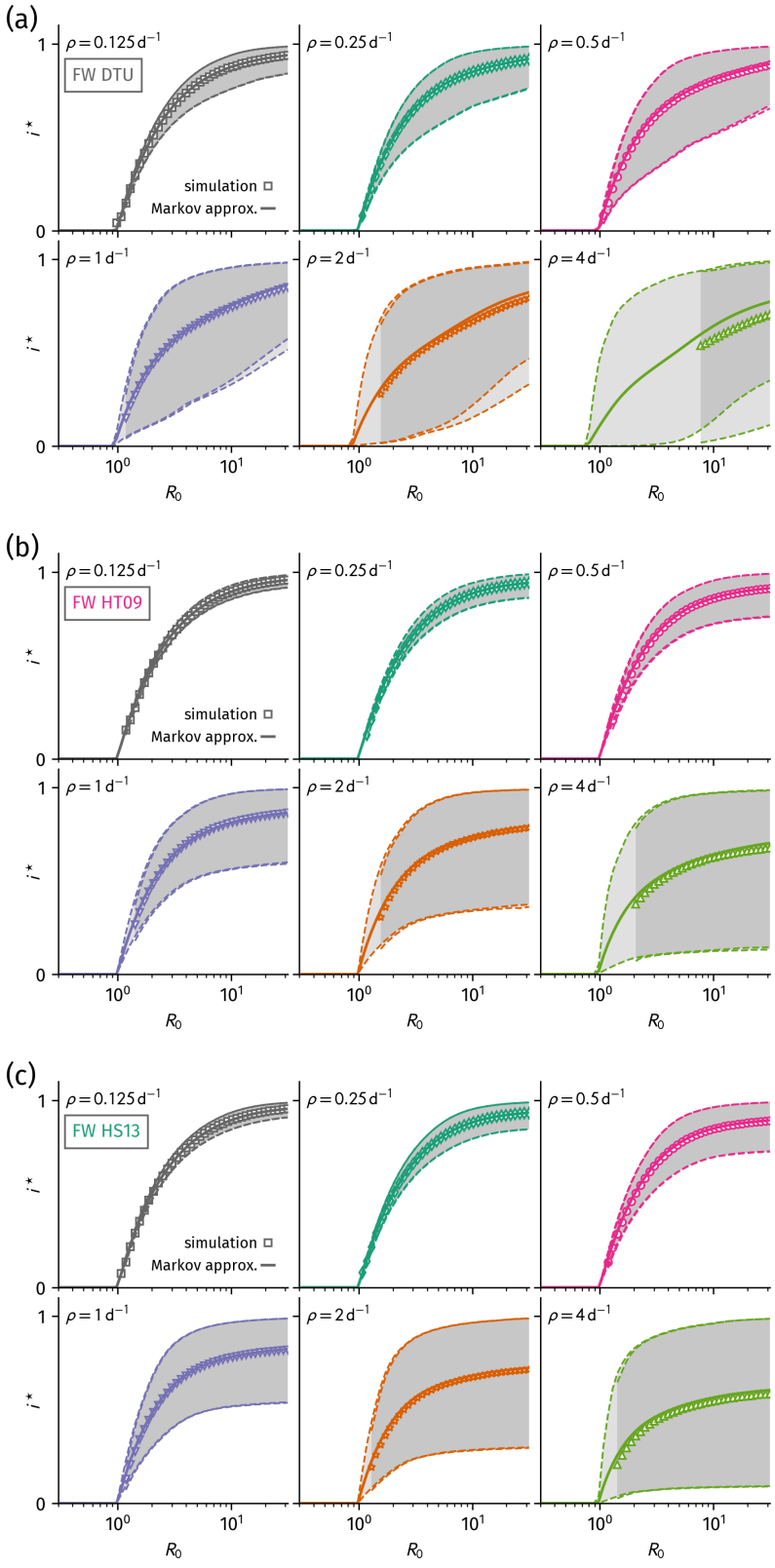


Figure E.6: The influence of circadian activity on epidemic response curves and epidemic thresholds, as a comparison between Gillespie simulations and the individual-based Markov-chain approximation. Since the networks investigated here become sparse and slowly varying at night (or during the weekend), the endemic state potentially decreases strongly during those times. Solid lines and markers represent the average endemic state Eq. (8.6) and grey bands mark the extent of their variation in the . When the minimal number of infected $i^\star_{\text{min}}$ (lower dashed lines) approaches $i^\star_{\text{min}} = 0$, the disease becomes unsustainable and dies out. The average endemic state is potentially much larger than the minimal endemic state and therefore, sudden jumps in the order parameter $i^\star$ occur.

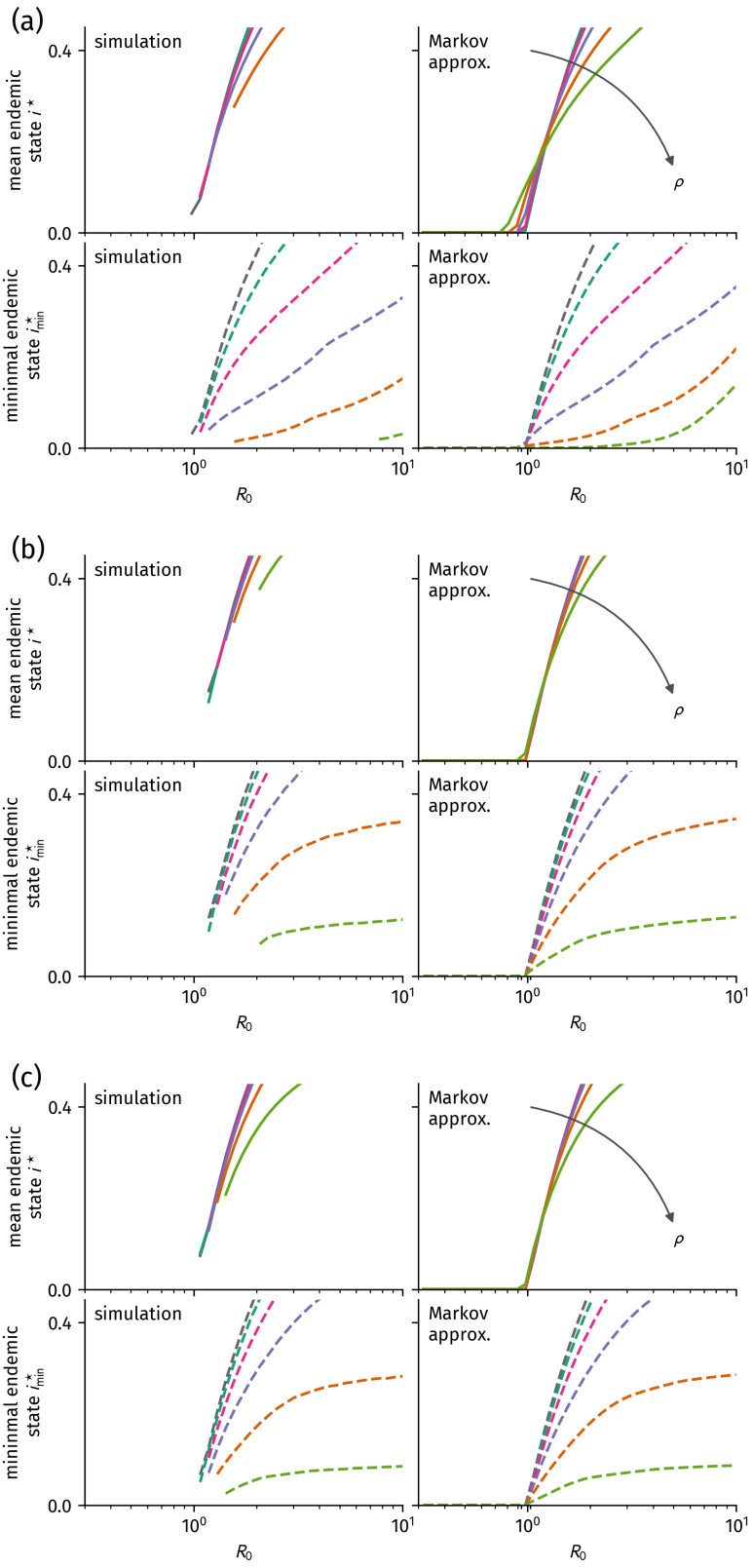


Figure E.7: Enlarged illustration of the influence of circadian activity on epidemic thresholds, as a comparison between Gillespie simulations and the individual-based Markov-chain approximation. The average minimal number of infected during an observation period is a more reliable order parameter to obtain the epidemic threshold. The Markov-chain approximation method gives increasingly differing predictions with increasing recovery rate ρ. (a) DTU Flockwork surrogate (FW surr.), (b) HT09 FW surr., (c) HS13 FW surr.

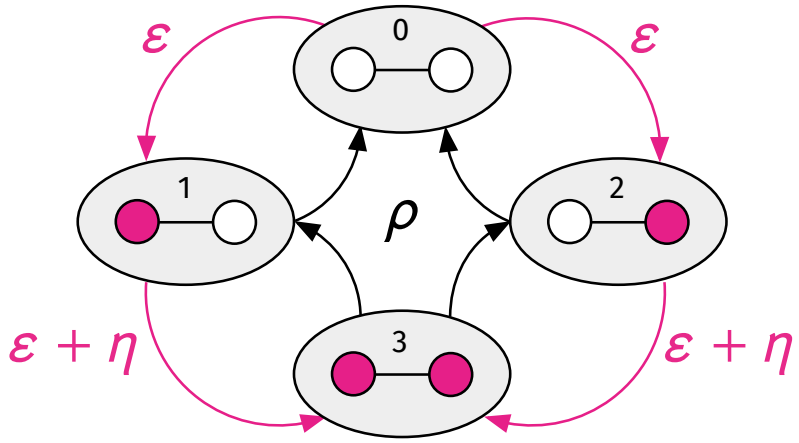


Figure E.8: The ε-SIS model on a single pair (complete network of $N = 2$).

$j_{\max} = 0$, one finds it as

$$R_0 = 1 + R_\epsilon. \tag{E.3}$$

Solving Eqs. (E.2) for the non-zero stationary state yields

$$\langle i_1 \rangle = \langle i_2 \rangle = \frac{R_0 + \sqrt{(1 - R_0 + R_\varepsilon)^2 + 4R_0R_\varepsilon} - R_\varepsilon - 1}{2R_0}.$$

The approximation made to obtain Eqs. (E.2) is that the normed covariance vanishes as $\varrho[i_1, i_2] = 0$, such that

$$\langle i_1 i_2 \rangle^{(\mathrm{MCA})} = \langle i_1 \rangle \langle i_2 \rangle .$$

E.5.2 Exact Evaluation

In a network consisting of a single pair, the two nodes can be in four states: none is infected (0), the first node is infected and the second node is healthy (1), the second node is infected and the first node is healthy (2), and both nodes are infected (3) (see Fig. E.8). Let w_j denote the probability that the system is in state j. The equations of evolution are given as

$$\partial_t \boldsymbol{w} = \boldsymbol{W} \cdot \boldsymbol{w} \tag{E.4}$$

with

$$\boldsymbol{W} = \begin{pmatrix} -2\varepsilon & \rho & \rho & 0 \\ \varepsilon & -\rho - \eta - \varepsilon & 0 & \rho \\ \varepsilon & 0 & -\rho - \eta - \varepsilon & \rho \\ 0 & \eta + \varepsilon & \eta + \varepsilon & -2\rho \end{pmatrix}.$$

Lets denote the basic reproduction number $R_0 = \eta/\rho$ and the spontaneous infection number $R_\varepsilon = \varepsilon/\rho$. The stationary state is found with $\partial_t \boldsymbol{w}^\star = 0$ which means it is given by the eigenvector of $\boldsymbol{W}$ corresponding to the eigenvalue $\omega = 0$, such that

$$\boldsymbol{w}^\star = C \begin{pmatrix} (R_\varepsilon(R_\varepsilon + R_0))^{-1} \\ (R_\varepsilon + R_0)^{-1} \\ (R_\varepsilon + R_0)^{-1} \\ 1 \end{pmatrix} \tag{E.5}$$

$$C = \left(1 + \frac{2}{R_\varepsilon + R_0} + \frac{1}{R_\varepsilon(R_\varepsilon + R_0)}\right)^{-1}.$$

where C is a normalization constant such that $\boldsymbol{w}^\star$ is a proper probability vector summing to $\sum_{j=0}^{3} w_j^\star = 1$. Note that the $\star$-superscript will be omitted in the following. The total number of infected in equilibrium is given as

$$\begin{aligned}\langle I \rangle &= (w_1 + w_2) + 2w_3 \\ &= 2C\left(1 + \frac{1}{R_\varepsilon + R_0}\right).\end{aligned}$$

The probability that a node is infected is

$$\begin{aligned}\langle i_1 \rangle &= w_1 + w_3 \\ \langle i_2 \rangle &= w_2 + w_3 = \langle i_1 \rangle\end{aligned}$$

and the probability that both nodes are infected is given as

$$\langle i_1 i_2 \rangle = w_3.$$

Hence, one finds the normed covariance

$$\begin{aligned}\varrho[i_1, i_2] &= \frac{\langle i_1 i_2 \rangle - \langle i_1 \rangle \langle i_2 \rangle}{\langle i_1 \rangle \langle i_2 \rangle} \\ &= \frac{R_0}{R_\varepsilon (1 + R_0 + R_\varepsilon)^2} > 0. \end{aligned} \tag{E.6}$$

A comparison with the MCA assumption $\varrho[i_1, i_2] = 0$ shows that the assumption is not justified.

E.5.3 Epidemic Threshold

Note that only in the limit $\varepsilon \to 0$ the ε-SIS-model becomes equal to the original SIS-model, which poses the following problem. Taking the limit $\varepsilon \to 0$ makes the disease-free state an absorbing state. Therefore, for any R_0, the single stationary state becomes equal to $w_0^\star = 1$ and $w_{j>0}^\star = 0$. It was shown in [142] that this is the case for any finite network: Only in thermodynamic equilibrium $N \to \infty$ there is non-zero probability to not be absorbed by the disease-free state. Essentially, this means that there is no epidemic threshold. However, this is only true for the strict definition of the epidemic threshold, which requires a non-linear system and that for values below the threshold, only the disease-free state exists, whereas above the threshold, two states exist: the endemic state and the absorbing disease-free state. In the following, two alternatives are proposed to find "an" epidemic threshold in a network of finite size.

On a Single Pair

In non-linear dynamics, an equilibrium is considered unstable if a small perturbation drives the system away from this equilibrium. Here, the "perturbation" will be a single infected. Using Eq. (E.4) and noting that the expected number of infected is given as $\langle I \rangle = w_1 + w_2 + 2w_3$, the temporal evolution of the total number of infected is

$$\partial_t \langle I \rangle = 2\varepsilon w_0 - \rho(w_1 + w_2 + 2w_3) + (\eta + \varepsilon)(w_1 + w_2).$$

When a single node is infected, the states (0) and (3) are unoccupied, such that $w_0 = w_3 = 0$ and $w_1 + w_2 = 1$. Then,

$$\partial_t \langle I \rangle = \eta - \rho - \varepsilon.$$

When this equation becomes positive, it means that the number of infected will grow, while for negative values, it will decay to a state of $\langle I \rangle < 1$ which means that the disease is not self-sustainable. The epidemic threshold is consequently found by setting $\partial_t \langle I \rangle = 0$. Introducing the basic reproduction number $R_0 = \eta/\rho$ and the spontaneous infection number $R_\varepsilon = \varepsilon/\rho$ yields

$$R_0 = 1 + R_\varepsilon,$$

which, surprisingly, is equal to the result of the MCA while the epidemic response curves completely differ.

The problem of the SIS-model always reaching the absorbing state is well-known and problematic for e.g. numerical simulations. To overcome this, an approach to compute the metastable (or quasi-stationary) state analytically is presented in the following.

One may assume that before being absorbed into the disease-free state, a metastable state is reached from which there is constant flux of probability into the disease-free state. The assumption is that the probability to find j infected nodes at time t is approximately

$$\pi_j(t) \approx \Big[1 - \pi_0(t)\Big] \pi_j^Q$$

where π_j^Q is the quasi-stationary probability that j nodes are infected and $\pi_0(t)$ is the probability that the system is in the disease-free state at time t. The probability π_j^Q is only defined for $j \geq 1$. In the ε-SIS-model on a pair, the quasi-stationary state is exact and given by the entries $w_{>0}$ in Eq. (E.5). Aggregating the states (1) and (2) yields the entries

$$\begin{aligned} \pi_1^Q &\propto 2(R_\varepsilon + R_0)^{-1} \\ \pi_2^Q &\propto 1. \end{aligned}$$

Further setting $R \equiv R_\varepsilon + R_0$ and renormalizing for the domain $j \geq 1$ one finds the quasi-stationary state

$$\begin{aligned} \pi_1^Q &= \frac{2}{R+2} \\ \pi_2^Q &= \frac{R}{R+2} \end{aligned}$$

with expected number of infected

$$\begin{aligned} \langle I \rangle^Q &= \pi_1^Q + 2\pi_2^Q \\ &= 2\frac{R+1}{R+2} \end{aligned}$$

and variance

$$\mathrm{Var}\,[I]^Q = \frac{2R}{(2+R)^2}.$$

Phenomenologically, these results should be equal to values found numerically when simulating with the quasi-stationary state algorithm [117] which similarly only samples from all states which are not the absorbing one. The results above come with a welcome convenience: when taking the limit $\varepsilon \to 0$, no divergencies arise and therefore the results should be correct for the original SIS-model.

One way to find epidemic thresholds from simulation data is to search for maxima in the noise. At a phase transition in finite systems, fluctuations in the order parameter are maximal (as opposed to diverging as they would in thermodynamic equilibrium). Typically, fluctuations are measured in two ways for SIS simulations, which are (a) the variance and (b) the variability/susceptibility

$$\chi = \frac{\mathrm{Var}\,[I]}{\langle I \rangle}.$$

Here, an additional measure for fluctuations will investigated, the (c) squared coefficient of variation

$$\mathrm{CV}[I]^2 = \frac{\mathrm{Var}\,[I]}{\langle I \rangle^2}.$$

For all three measures, the maximum is easily found by differentiation:

$$\begin{aligned} R^{\mathrm{Var}} &= 2 \\ R^{\chi} &= \sqrt{2} \\ R^{\mathrm{CV}} &= 1. \end{aligned}$$

Comparing with the results of the ODE analysis and the MCA, the closest consistent result is the one obtained from maximum in the coefficient of variation where

$$R_0 = 1 - R_\varepsilon.$$

One may further compute the expected node status and correlation as

$$\begin{aligned} \langle i_1 \rangle^Q &= \frac{R+1}{R+2} \\ \langle i_1 i_2 \rangle^Q &= \frac{R}{R+2} \end{aligned}$$

such that even in the quasi-stationary state of the original SIS-model, one finds non-zero normalized covariance

$$\varrho[i_1, i_2]^Q = -\frac{1}{(R+1)^2}.$$

On Complete Networks

The stationary state of the ε-SIS system on any N-sized complete network was derived in [142] to be

$$\pi_{j>0} = \pi_0 \binom{N}{j} R_\varepsilon \left(\frac{R_0}{N-1}\right)^{j-1} \frac{\Gamma(R_\varepsilon(N-1)/R_0 + j)}{\Gamma(R_\varepsilon(N-1)/R_0 + 1)}.$$

where $R_\varepsilon = \varepsilon/\rho$ and $R_0 = \eta(N-1)/\rho$. Therefore, the quasi–stationary state scales as

$$\pi_j^Q \propto \left(\begin{array}{c} N \\ j \end{array} \right) \left(\frac{R_0}{N-1} \right)^{j-1} \Gamma(R_\varepsilon(N-1)/R_0 + j).$$

After taking the limit $\varepsilon \to 0$, it is found as

$$\pi_j^Q = C \frac{N!}{(N-j)!} \frac{1}{j} \left(\frac{R_0}{N-1} \right)^{j-1}$$
$$C = \sum_{j=1}^{N} \frac{N!}{(N-j)!} \frac{1}{j} \left(\frac{R_0}{N-1} \right)^{j-1}.$$

From which all other observables can be found analogously to the pair network. Using $\left(\begin{array}{c} N \\ j \end{array} \right) \approx \frac{N^j}{j!}$ for $j \ll N$, one further sees that at $R_0 = 1$, the probability to be in a state of j infected is

$$\pi_j^Q(R_0 = 1) \propto \frac{1}{j}$$

which suggests that $R_0 = 1$ is the epidemic threshold for $N \gg 1$, since at phase transitions, the order parameter is distributed according to a power-law.

F

Rate and State-Space Trajectory Inference from Real Data

In Sec. 7.2, a method to infer the state-space trajectory of time-varying rate Flockwork simulations is introduced and applied to both artificial Flockworks and real-world data. The full analyses for increasing Δt are shown in the following.

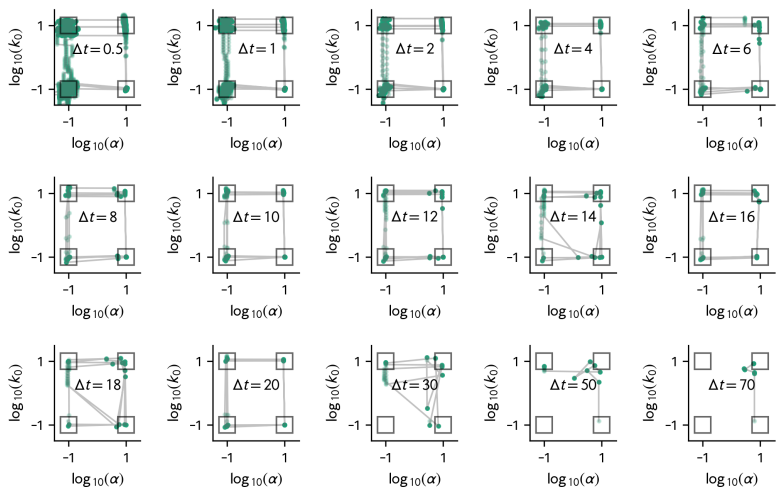


Figure F.1: Inference of the (α, k_0)-state-space trajectory from a single Flockwork simulation as described in Sec. C.4 for increasing inference parameter Δt.

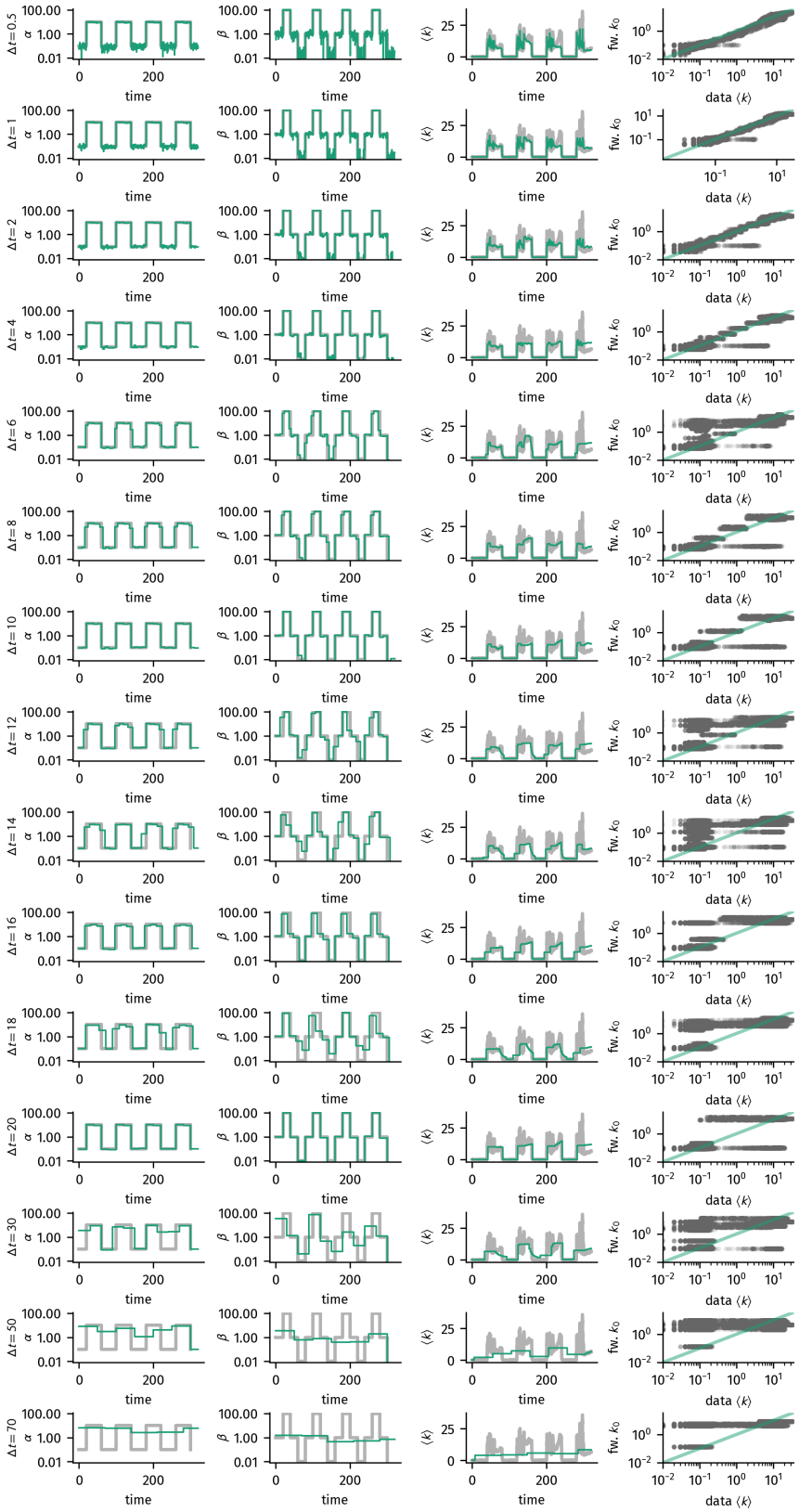


Figure F.2: Inference of the reconnection rates α and β from a single Flockwork simulation as described in Sec. C.4 for increasing inference parameter Δt.

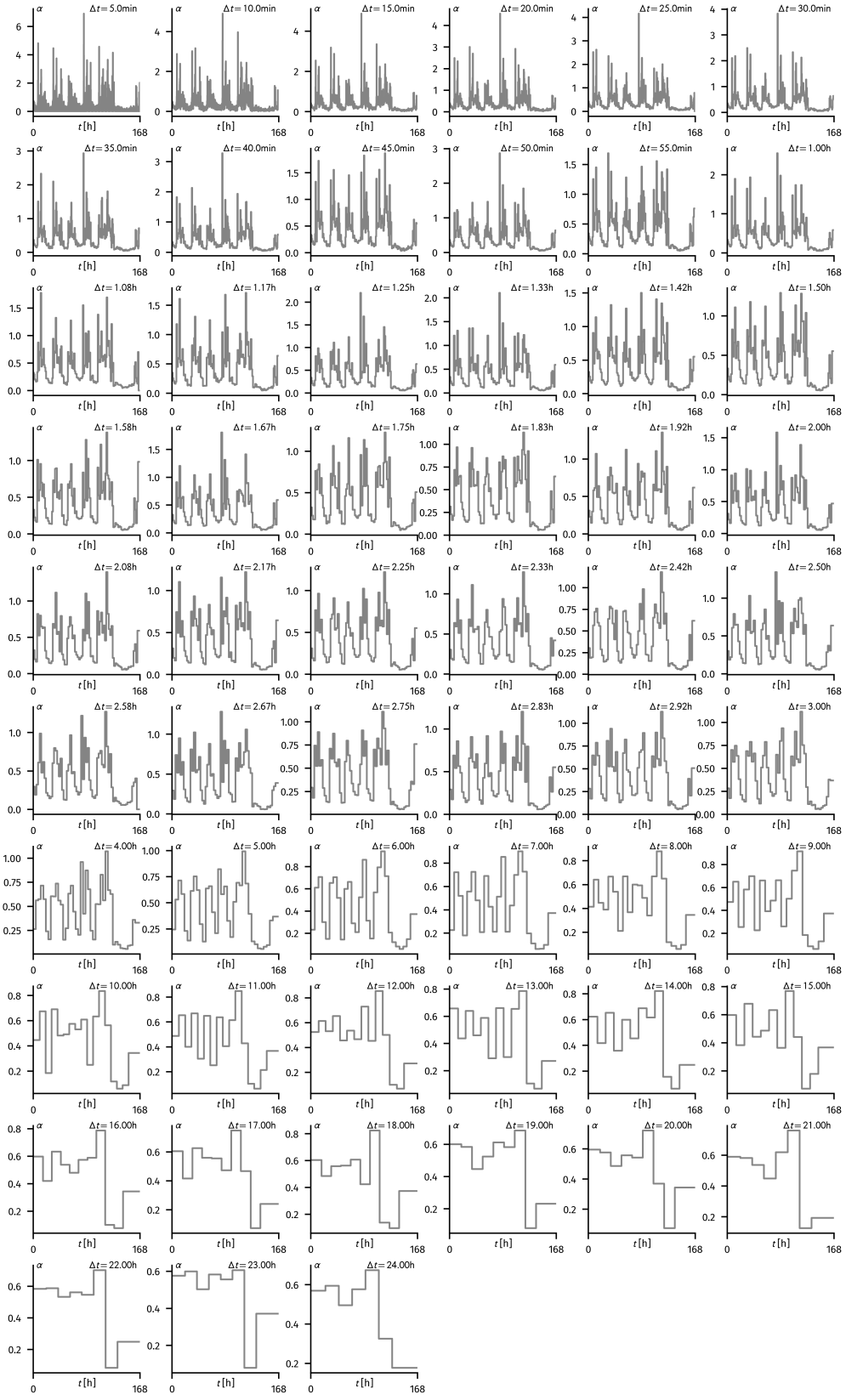


Figure F.3: Inference of the reconnection rate α from the DTU data for increasing inference parameter Δt.

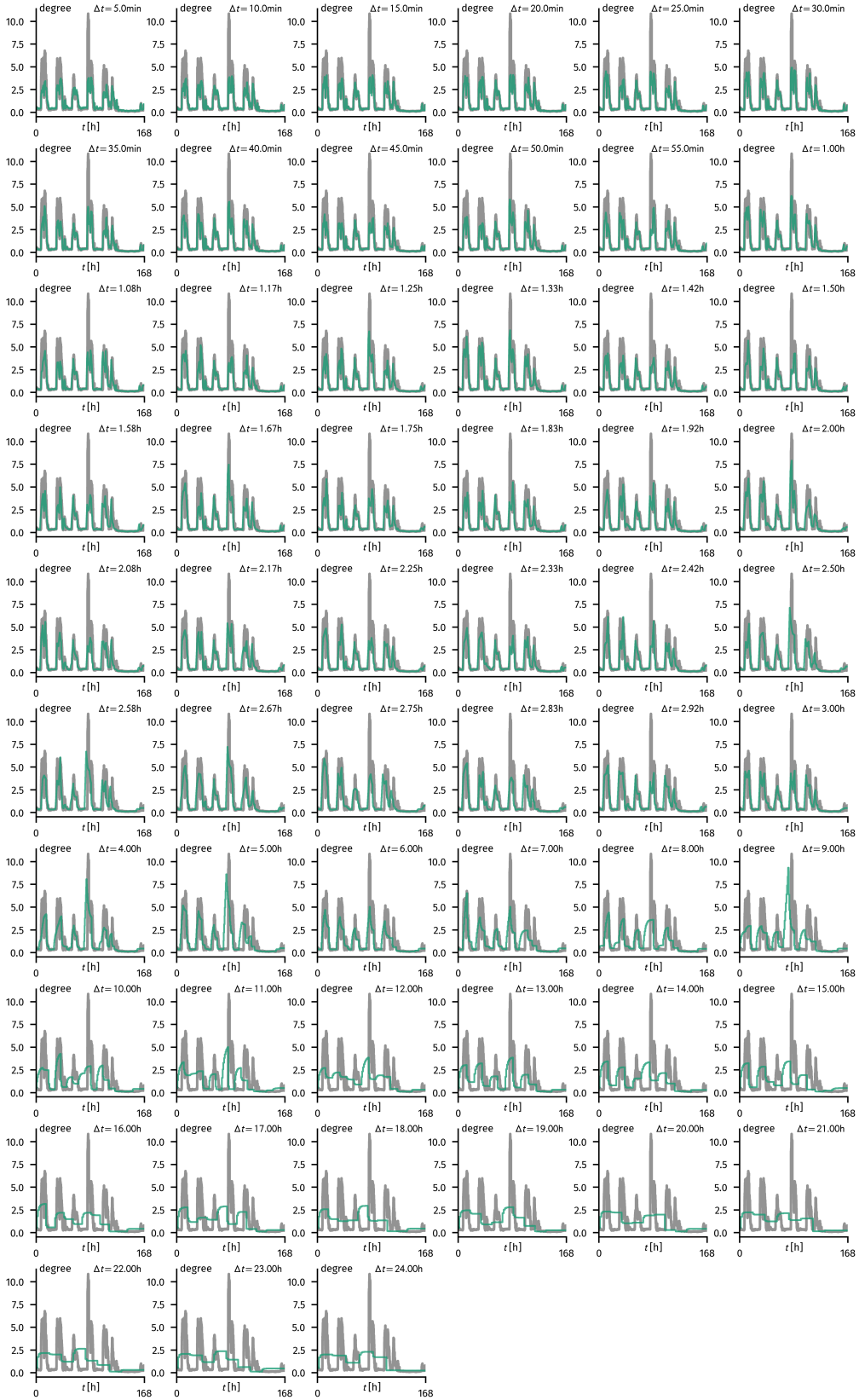


Figure F.4: Inference of the expected mean degree k_0 from the DTU data for increasing inference parameter Δt and comparison to the actual mean degree $\langle k \rangle (t)$.

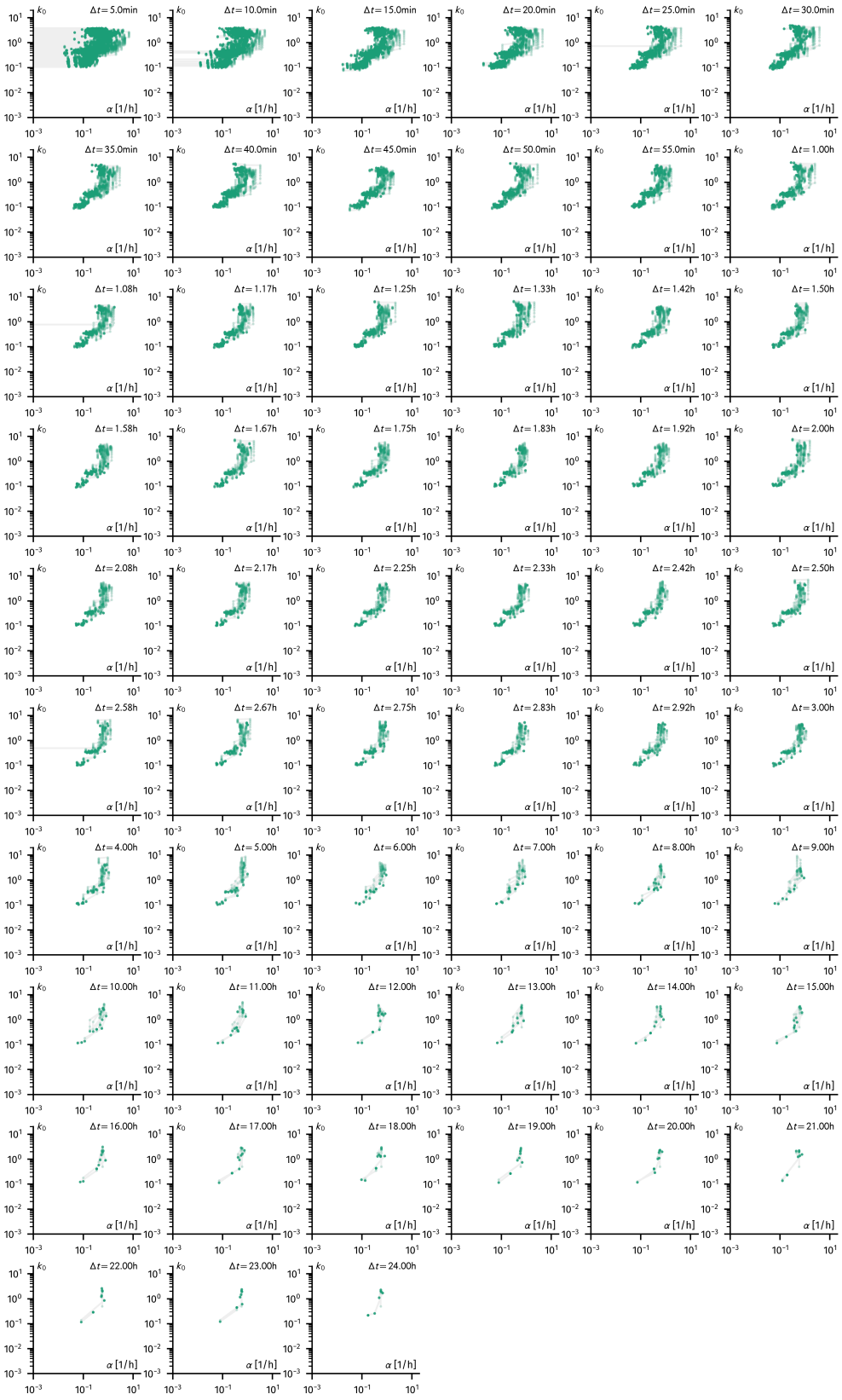


Figure F.5: Inferred trajectory in the (α, k_0)-plane from the DTU data for increasing inference parameter Δt

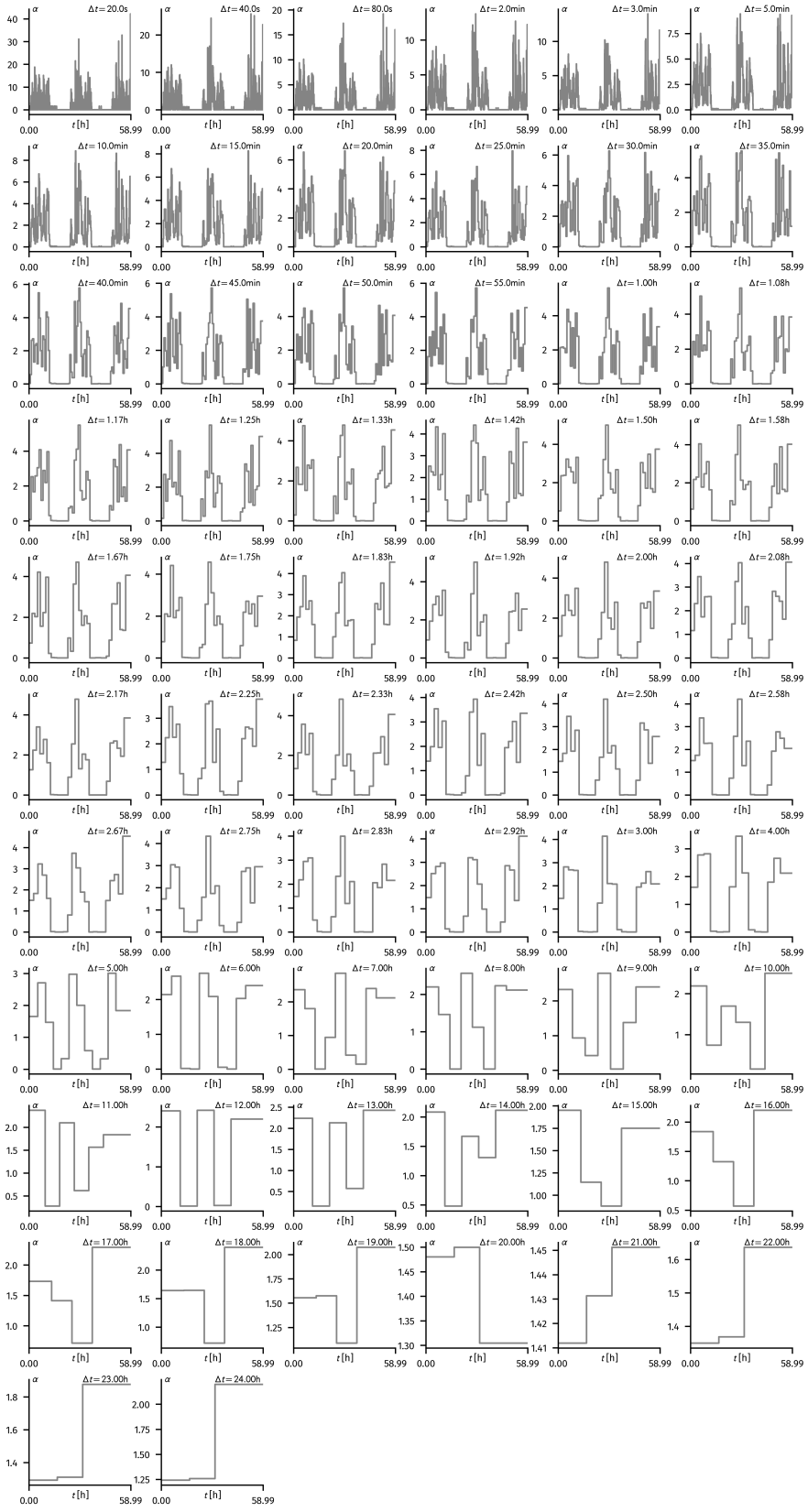


Figure F.6: Inference of the reconnection rate α from the HT09 data for increasing inference parameter Δt.

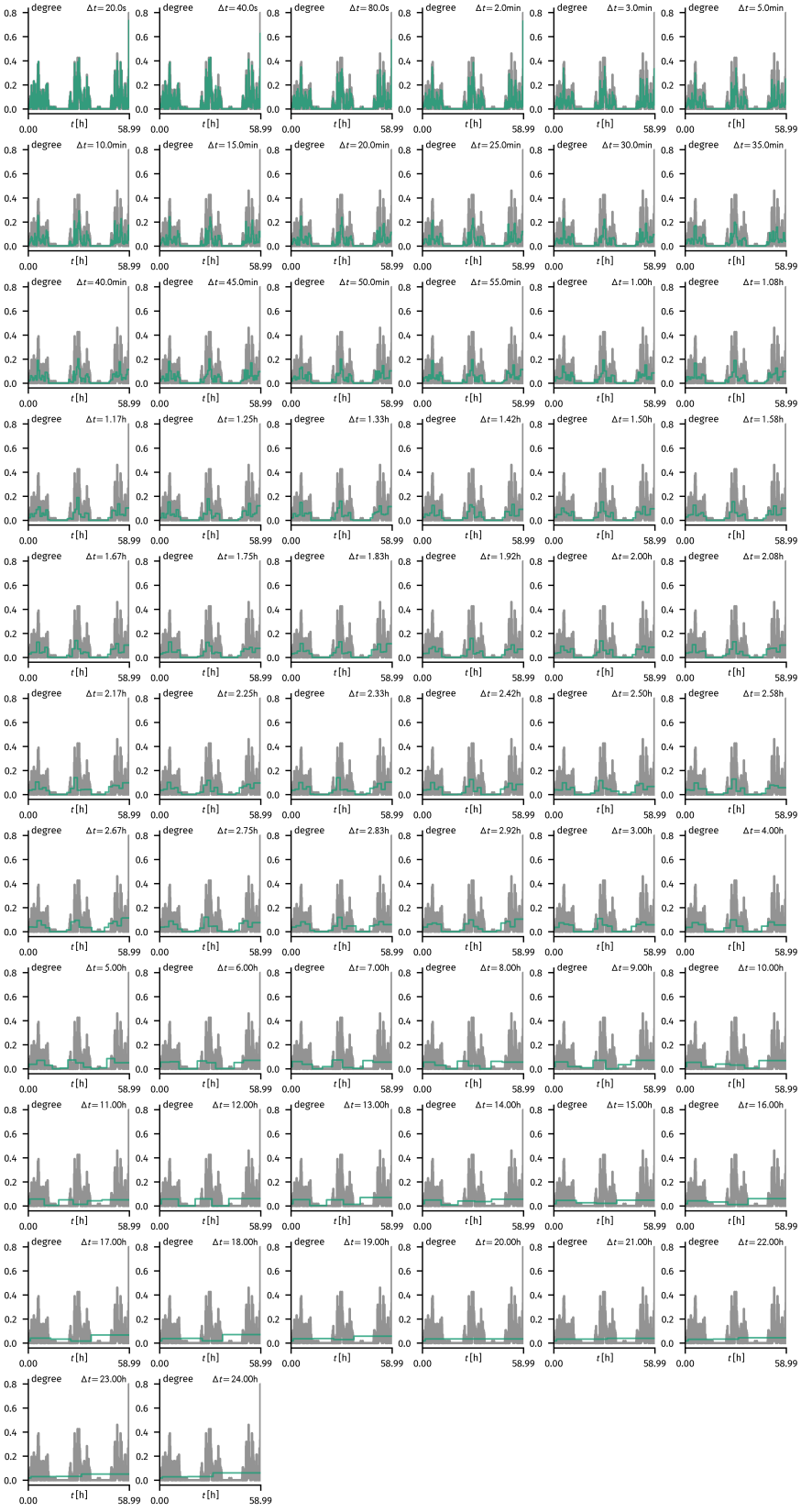


Figure F.7: Inference of the expected mean degree k_0 from the HT09 data for increasing inference parameter Δt and comparison to the actual mean degree $\langle k \rangle (t)$.

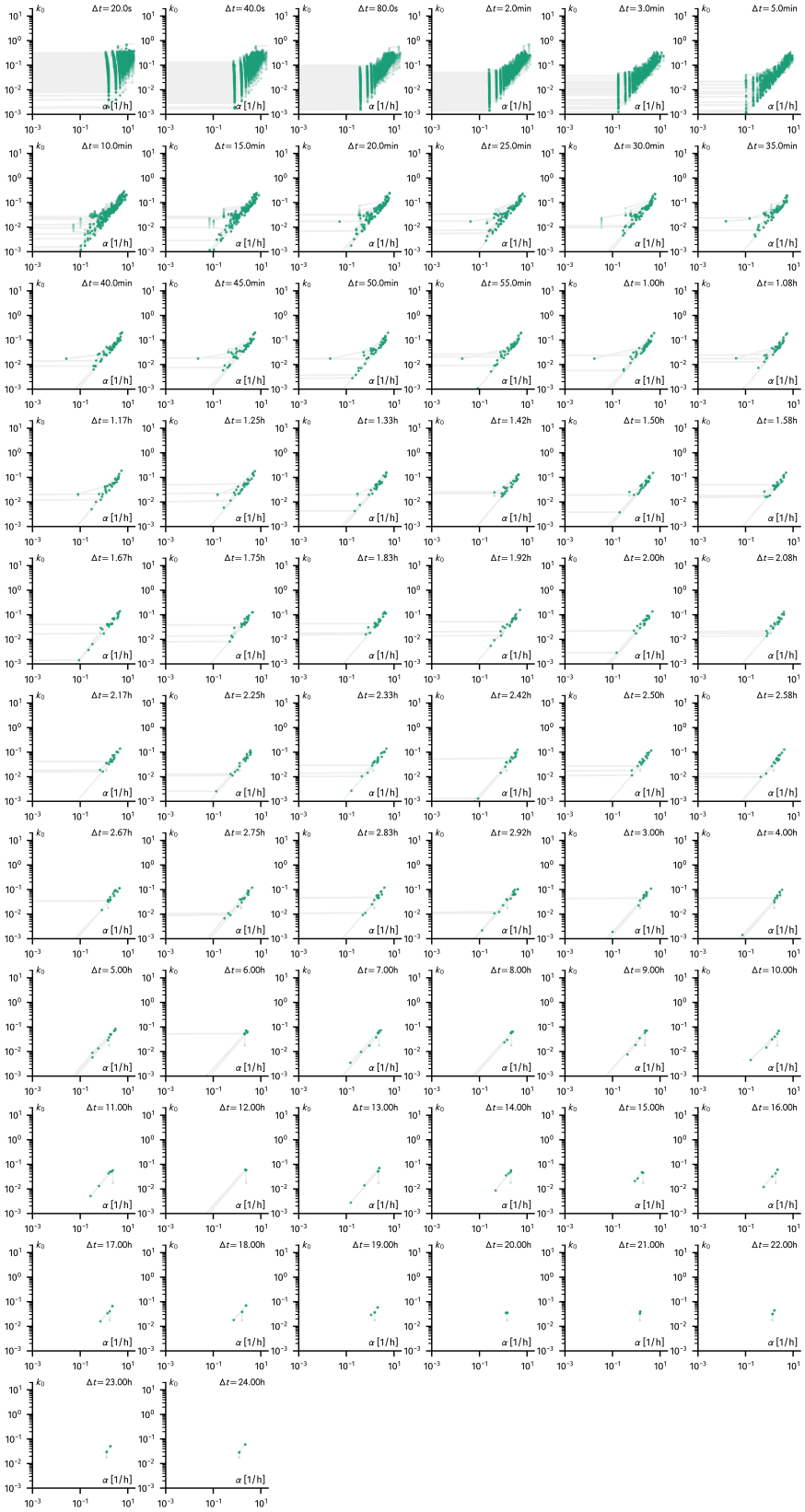


Figure F.8: Inferred trajectory in the (α, k_0)-plane from the HT09 data for increasing inference parameter Δt

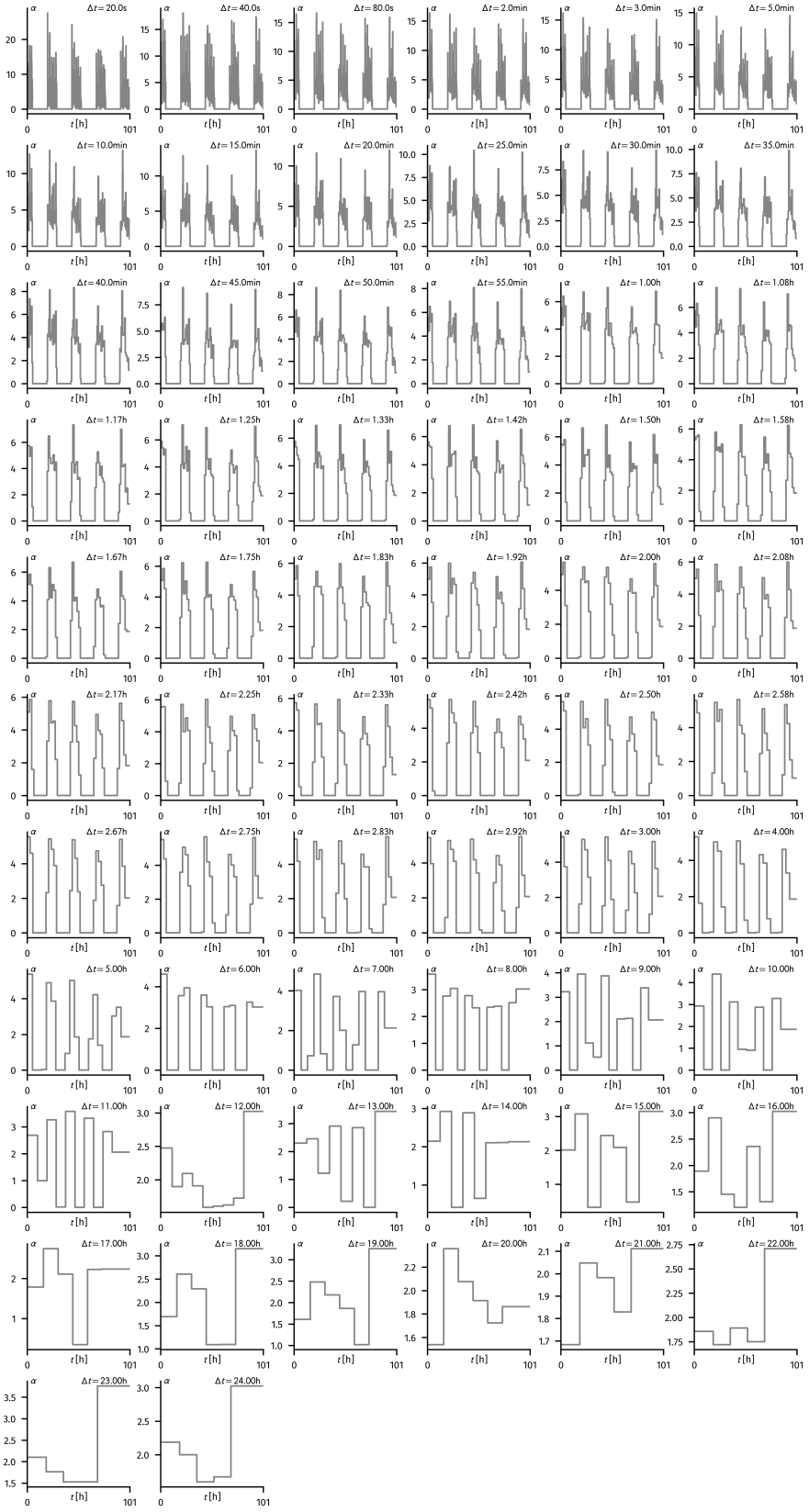


Figure F.9: Inference of the reconnection rate α from the HS13 data for increasing inference parameter Δt.

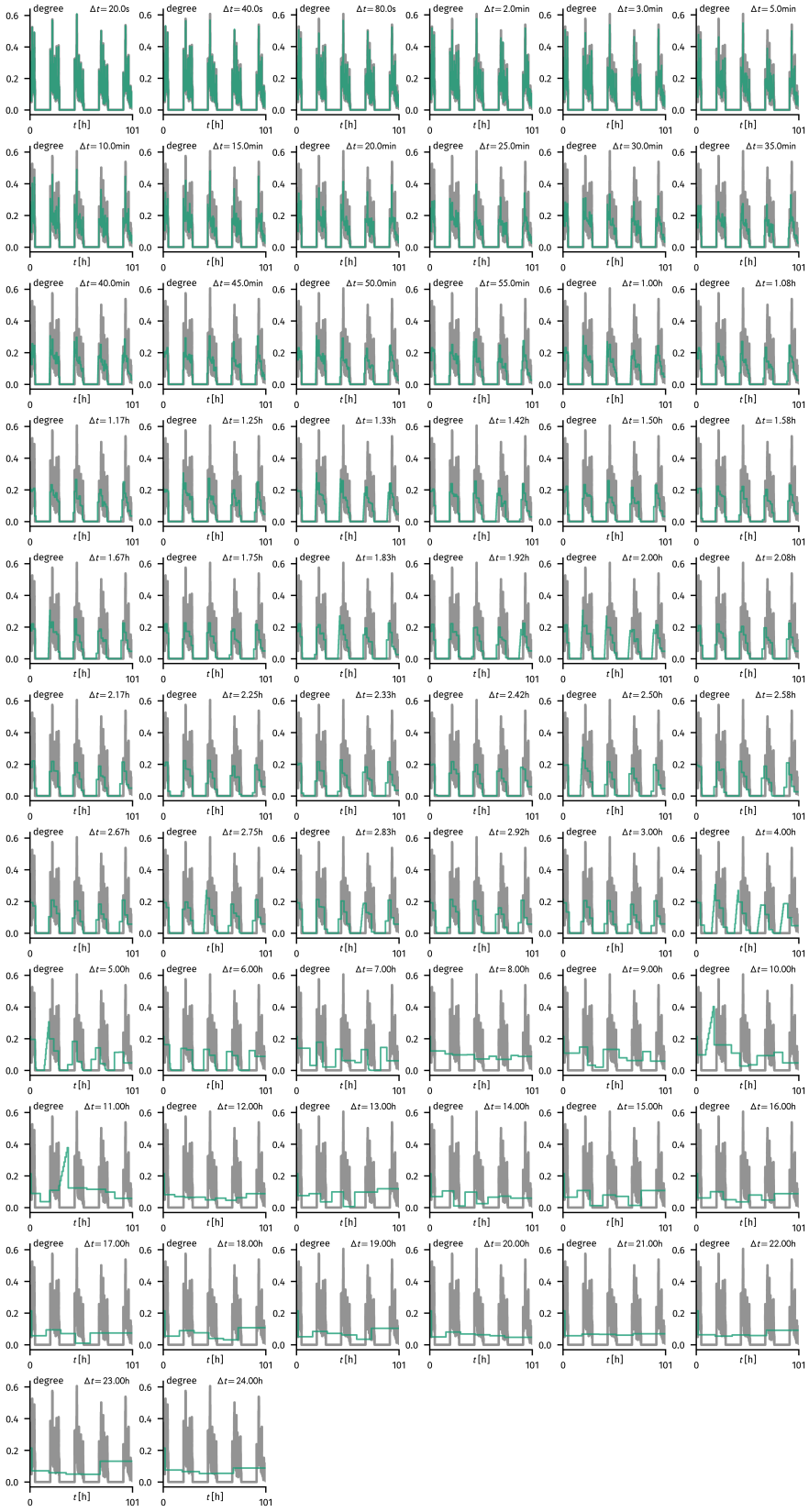


Figure F.10: Inference of the expected mean degree k_0 from the HS13 data for increasing inference parameter Δt and comparison to the actual mean degree $\langle k \rangle (t)$.

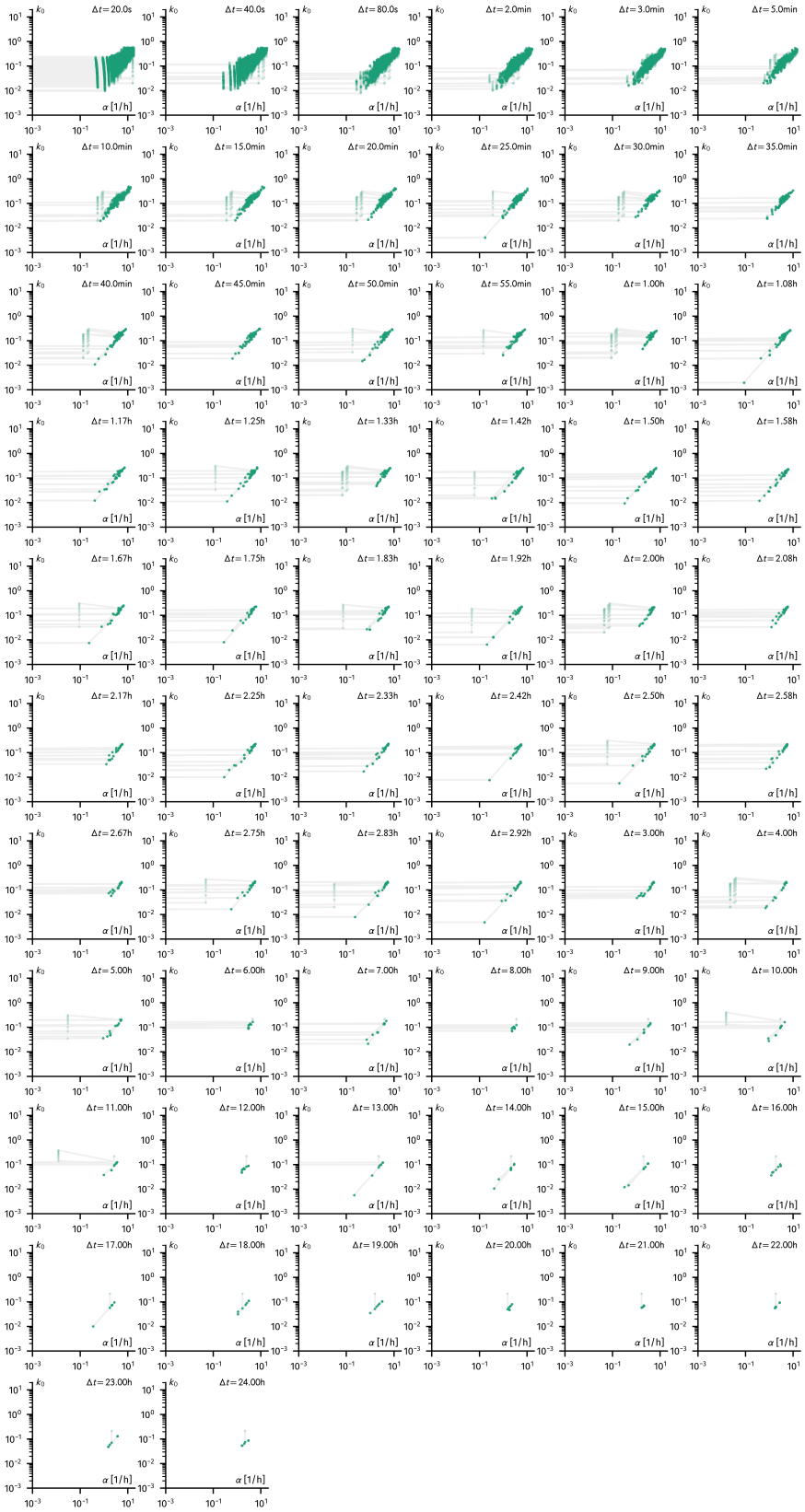


Figure F.11: Inferred trajectory in the (α, k_0)-plane from the HS13 data for increasing inference parameter Δt

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig ohne fremde Hilfe verfasst und nur die angegebene Literatur und Hilfsmittel verwendet zu haben.

Ort Datum

Benjamin F. Maier